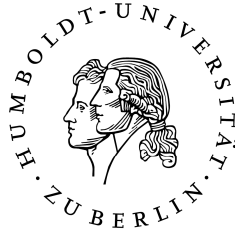


HUMBOLDT-UNIVERSITÄT ZU BERLIN

INSTITUT FÜR BIBLIOTHEKS- UND INFORMATIONSWISSENSCHAFT



BERLINER HANDREICHUNGEN
ZUR BIBLIOTHEKS- UND
INFORMATIONSWISSENSCHAFT

HEFT 468

EPISTEMISCHE FUNKTIONEN VON ZITIERUNGEN
EINE VERGLEICHENDE ANALYSE VON ZITATIONSFUNKTIONEN UND
DEREN IMPLIKATIONEN FÜR DIE MODELLIERUNG VON
ZITATIONSNETZWERKEN

VON
CARA LUZIA SEITZ

EPISTEMISCHE FUNKTIONEN VON ZITIERUNGEN
EINE VERGLEICHENDE ANALYSE VON ZITATIONSFUNKTIONEN UND
DEREN IMPLIKATIONEN FÜR DIE MODELLIERUNG VON
ZITATIONSNETZWERKEN

VON
CARA LUZIA SEITZ

Berliner Handreichungen zur
Bibliotheks- und Informationswissenschaft

Begründet von Peter Zahn
Herausgegeben von
Vivien Petras
Humboldt-Universität zu Berlin

Heft 468

Seitz, Cara Luzia

Epistemische Funktionen von Zitierungen : Eine vergleichende Analyse von Zitationsfunktionen und deren Implikationen für die Modellierung von Zitationsnetzwerken / von Cara Luzia Seitz. - Berlin : Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin, 2021. – 81 S. : graph. Darst. - (Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft ; 468)

ISSN 14 38-76 62

Abstract

In bibliometrischen Analysen werden häufig Zitationsbeziehungen dazu genutzt thematische Cluster in Forschungsfeldern zu identifizieren. Aufgrund mangelnder Informationen wird dabei nicht zwischen verschiedenen thematischen Bezügen der Zitierungen unterschieden, also ob sich ein Zitat auf etwa eine Theorie, eine Methode, einen Forschungsgegenstand oder anderes bezieht. In thematischen Mappings vermischen sich diese Bezüge und sind nicht mehr zu unterscheiden. Es wurde bisher nicht systematisch untersucht, ob und wie diese Vermischung das Mapping beeinflusst. Die vorliegende Arbeit beschäftigt sich mit (1) der Analyse epistemischer Funktionen von Zitationsbeziehungen und (2) deren Auswirkungen auf das bibliometrische Mapping thematischer Strukturen. Dazu werden im ersten Teil der Analyse die Referenzen einer begrenzten Anzahl an Publikationen nach ihrer Funktion kodiert und ausgewertet und im zweiten Teil der Analyse Netzwerkkarten unter Berücksichtigung der Zitationsfunktionen erstellt und analysiert. Dabei werden Methoden der qualitativen Inhaltsanalyse und netzwerkanalytische Verfahren eingesetzt.

Diese Veröffentlichung geht zurück auf eine Masterarbeit im weiterbildenden Masterstudiengang im Fernstudium Bibliotheks- und Informationswissenschaft (Library and Information Science, M. A. (LIS)) an der Humboldt- Universität zu Berlin.

Eine Online-Version ist auf dem edoc Publikationsserver der Humboldt-Universität zu Berlin verfügbar.



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung - Nicht kommerziell - Keine Bearbeitungen 4.0 International Lizenz.

Inhaltsverzeichnis

1. Einleitung	9
2. Theoretische Einordnung und Stand der Forschung	12
2.1. Theoretische Einordnung	12
2.2. Forschungsstand thematisches Mapping	14
2.3. Forschungsstand Zitationsfunktionen	17
3. Methodisches Vorgehen	21
3.1. Sampling	21
3.1.1. Sampling Phase 1 – Klassifizierung von Zitationsfunktionen	23
3.1.2. Sampling Phase 2 – Netzwerkanalyse	25
3.2. Klassifikation	27
3.3. Netzwerkmodellierung	32
4. Ergebnisse	34
4.1. Analyse der Klassifizierung	34
4.1.1. Überblick Datensatz und kodierte Segmente	34
4.1.2. Epistemische Funktion	35
4.1.3. Ort der Zitierung	36
4.1.4. Selbst-Zitation	39
4.1.5. Zuordnung Felddatensatz und Clusterzurodnung	40
4.1.6. Erscheinungsjahre der zitierten Quellen	41
4.2. Analyse der Netzwerke	43
4.2.1. Überblick	43
4.2.2. Epistemische Funktion: theoretical	46
4.2.3. Epistemische Funktion: method	48
4.2.4. Epistemische Funktion: empirical object	50
4.2.5. Epistemische Funktion: research problem	52
4.2.6. Zusammenfassung	55
5. Diskussion	57
6. Literaturverzeichnis	61

A. Anhang – Dokumentation und Material zum Sampling	66
A.1. Dokumentation Lexical Query Felddatensatz Invasionsbiologie	66
A.2. Übersicht Sample Invasionsbiologie	71
A.2.1. Phase 1: Analyse Klassifikation von Zitationsfunktionen	71
A.2.2. Phase 2: Netzwerkanalyse	75
B. Anhang – weitere Tabellen und Abbildungen zur Auswertung der Ergebnisse	81
B.1. Tabellen	81

Abbildungsverzeichnis

3.1. Zuordnung der zufällig gezogenen Publikationen hinsichtlich Invasionsbezug	25
3.2. Zuordnung der invasionsbiologischen Publikationen hinsichtlich Forschungsansatz . .	25
3.3. Screenshot Kodierung mit dem Datenanalyseprogramm MaxQDA	30
3.4. Screenshot Kodierung und Dokumentmemos mit dem Datenanalyseprogramm MaxQDA	33
4.1. Prozentuale Verteilung der Orte der Zitierungen nach epistemischer Funktion	37
4.2. Verteilung der Zitate nach standardisierten Kapiteln	38
4.3. Prozentuale Verteilung der Zitate nach standardisierten Kapiteln für die drei Doku- mentgruppen	39
4.4. Verteilung der keinem Cluster zuordenbaren Zitationen nach epistemischer Funktion in %	41
4.5. Zuordnung zu Clustern im Felddatensatz nach epistemischer Funktion der zitierten Quellen	41
4.6. Darstellung der Verteilungen von Publikationsjahren nach epistemischer Funktion (Box- plots)	42
4.7. Primärnetzwerk mit Hervorhebung der in den Teilnetzwerken enthaltenen Publikatio- nen und Verbindungen	44
4.8. Primärnetzwerk mit Hervorhebung der Anteile eingehender Verbindungen nach episte- mischen Funktion	45
4.9. Teilnetzwerk „theoretical“	47
4.10. Teilnetzwerk „method“	49
4.11. Teilnetzwerk „empirical object“, Topics 1: Name	51
4.12. Teilnetzwerk „empirical object“, Topics 2: Eigenschaft	52
4.13. Teilnetzwerk „research problem“	54

Tabellenverzeichnis

3.1. Übersicht Abfrage WoS-IDs	22
3.2. Übersicht und Kennzeichen der verschiedenen Forschungsansätze in der Invasionsbiologie	24
4.1. Übersicht Auswertung Fallzahlen	34
4.2. Übersicht absolute Häufigkeiten: epistemic_function	35
4.3. Übersicht absolute Häufigkeiten: location_stand	36
4.4. Übersicht absolute Häufigkeiten: self_citation	39
4.5. Übersicht Angaben Felddaten	40
4.6. Übersicht Topics „research problem“ (häufigste 82%)	53

1. Einleitung

Die Produktion und Vermehrung wissenschaftlichen Wissens vollzieht sich auf den Schultern von Riesen. In diesem Sinne bauen neue wissenschaftliche Erkenntnisse immer auf ihnen vorausgehenden Arbeiten auf. Diese Einsicht schlägt sich auch in der Tradition der Zitierung relevanter Arbeiten und Erkenntnisse in wissenschaftlichen Publikationen nieder. Durch unterschiedliche Beweggründe motiviert und auf verschiedene epistemische Inhalte bezogen, können Zitationen in wissenschaftlichen Texten von ganz verschiedener Art sein. Mit diesen verschiedenen epistemischen Funktionen von Zitierungen beschäftigt sich die vorliegende Arbeit.

In der Wissenschaftsforschung hat es sich durchgesetzt die Formen und Mechanismen wissenschaftlicher Erkenntnisproduktion innerhalb einzelner Felder zu erforschen, da sich zwischen den Feldern zum Teil sehr unterschiedliche Kulturen entwickelt haben. Als Teilbereich der Wissenschaftsforschung leistet die Bibliometrie mit ihren spezifischen methodischen Ansätzen einen besonderen Beitrag auf diesem Gebiet. Die Bibliometrie hat Methoden entwickelt wissenschaftliche Felder voneinander abzugrenzen und thematische Strukturen zu identifizieren. Dazu werden die in bibliographischen Daten vorhandenen Zitationsbeziehungen benutzt, um thematisch naheliegende Publikationen zu identifizieren und nach Forschungsthemen zu clustern. Die bibliographischen Daten für diese Analysen stammen meist aus großen, kommerziellen Datenbanken wie Web of Science oder Scopus und bieten eine eingeschränkte Informationsgrundlage. Es werden beispielsweise keine Informationen darüber bereitgestellt welchen thematischen Bezug eine Zitation herstellt oder an welcher Stelle die Zitation im Text auftritt. Tatsächlich drücken sich in Zitationsbeziehungen aber sehr unterschiedliche thematische Bezüge aus. So ist es von inhaltlicher Bedeutung, ob eine Zitation etwa auf die Verwendung einer bestimmten Methode verweist, die Forschungsarbeit in einem bestimmten theoretischen Rahmen einordnet, sich auf Informationen über den Untersuchungsgegenstand bezieht oder die Forschungsfrage motiviert. Da in den üblicherweise verwendeten bibliographischen Daten keine Informationen zu diesen verschiedenen thematischen Perspektiven von Zitationsbeziehungen enthalten sind, bleiben diese in bisherigen Analysen stets unberücksichtigt; ein Umstand der die thematische Strukturierung der produzierten Karten in bisher nicht systematisch untersuchter Art und Weise beeinflusst.

Erklärtes Ziel der vorliegenden Arbeit ist an diese Überlegungen anschließend (1) die Entwicklung und Dokumentation einer Vorgehensweise zur Berücksichtigung von Zitationsfunktionen bei der Analyse thematischer Strukturen, (2) die vergleichende Analyse der epistemischen Funktionen von Zitationsbeziehungen in einem ausgewählten wissenschaftlichen Feld und (3) die Exploration der Implikationen, die eine Berücksichtigung bzw. Nicht-Berücksichtigung von Zitationsfunktionen auf das bibliometrische Mapping thematischer Strukturen hat.

Damit wird zum einen eine erste Einschätzung der konzeptionellen und praktischen Realisierbarkeit der Klassifikation von Zitierungen nach ihren epistemischen Dimensionen ermöglicht, zum anderen werden empirische Einblicke in die Unterscheidbarkeit von Publikationen und ihnen zugrundeliegenden Forschungsansätzen hinsichtlich der Struktur ihrer Zitierungen gegeben. Die Analyse auf Grundlage verschiedener epistemischer Funktionen verspricht Einsichten in die thematische Mehrdimensionalität

von Zitationsnetzwerken und ermöglicht die Erörterung eines Informationszugewinns für die Analyse thematischer Strukturen in Zitationsnetzwerken.

Die vorliegende Arbeit schließt sich theoretisch an Arbeiten aus dem Bereich Wissenschaftsforschung an. Theoretisch lässt sie sich in den Kontext einiger wissenschaftssoziologischer, bedeutender Arbeiten stellen, die in Abschnitt 2.1 kurz angerissen werden. Dadurch soll die Perspektive der Arbeit und die Einbettung der Fragestellung in übergeordnete Kontexte verdeutlicht werden. Forschungspraktisch bezieht die Arbeit einige relevante Vorarbeiten ein. Der Forschungsstand zum bibliometrischen Mapping thematischer Strukturen wird daher in Abschnitt 2.2 insbesondere unter Berücksichtigung verschiedener methodischer Vorgehensweisen zusammengefasst. In Abschnitt 2.3 werden einige frühere Arbeiten zu den verschiedenen Funktionen von Zitierungen dargestellt. Schlaglichter werden dabei vor allem auf die verschiedenen methodischen Vorgehensweisen und die Ausarbeitung unterschiedlicher Kategoriensysteme zur Klassifikation von Zitationsfunktionen geworfen. Durch die Reflektion bisheriger Forschungsarbeiten kann gezeigt werden, dass das Mapping thematischer Strukturen auf Grundlage bibliometrischer Daten zwar ein viel bearbeitetes Forschungsgebiet darstellt, die systematische Berücksichtigung von Zitationsfunktionen dabei aber bisher unberücksichtigt bleibt.

Das Forschungsdesign gliedert sich in zwei Phasen. Die erste Phase widmet sich der Analyse der epistemischen Funktionen von Zitationsbeziehungen. Die zweite Phase widmet sich der Darstellung und Analyse von Zitationsnetzwerken unter Berücksichtigung ihrer epistemischen Funktionen. Datengrundlage für das empirische Vorgehen beider Phasen bildet der „Felddatensatz Invasionsbiologie“ der im Rahmen des Forschungsprojekts „Fachspezifische Formen von Open Science“ unter der Leitung von Dr. Theresa Velden am Deutschen Zentrum für Hochschul- und Wissenschaftsforschung erhoben wurde. Der Felddatensatz für die Invasionsbiologie wurde mittels einer lexikalischer Abfrage in Web of Science erzeugt. Berücksichtigung fanden Publikationen, die innerhalb des Zeitfensters 2000–2019 publiziert wurden und im Science Citation Index Expanded oder im Social Sciences Citation Index enthalten sind. Für die Klassifizierung von Zitationsfunktionen (Phase eins) werden neun Publikationen aus der Grundgesamtheit des Felddatensatzes ausgewählt. Das Sampling folgt dabei einem mehrstufigen Stichprobenverfahren, welches sowohl Zufallselemente aufweist, als auch mit bewussten, theoretisch geleiteten, Auswahlentscheidungen vorgeht. Da Forschungsprozesse und -praktiken innerhalb ein und desselben wissenschaftlichen Spezialgebietes zum Teil stark variieren, wird die feldinterne Heterogenität der Forschungsansätze bei der Auswahl des Samples berücksichtigt. Für die netzwerkanalytische Auswertung (Phase zwei) wird eine Publikation aus der Grundgesamtheit des Felddatensatzes ausgewählt. Diese Publikation dient als Seed Dokument von dem ausgehend sowohl die ihn zitierenden als auch die durch ihn zitierten Referenzen als Grundlage für die Erstellung der Netzwerke dienen. Das mehrstufige Auswahlverfahren des Seed Dokuments folgt dabei einigen theoretischen Überlegungen, beinhaltet aber auch Elemente einer Zufallsauswahl. Alle Schritte und Überlegungen zum Sampling sind in Abschnitt 3.1 ausführlich dokumentiert.

Für beide Phasen relevant ist die Einschätzung der Zitationsbeziehungen hinsichtlich ihrer epistemischen Funktion. Für die vergleichende Analyse wissenschaftlicher Felder als epistemische Regime schlagen Gläser u. a. (2018) vor, sie hinsichtlich der verwendeten Methoden, der untersuchten Gegenstände, des bearbeiteten Forschungsproblems, der Bezugnahme auf praktische Anwendungskontexte und der theoretischen Ausrichtung zu unterscheiden und vergleichen. Aus diesen konkreten Überlegungen wurden die fünf Kategorien für die Unterscheidung von Zitationsfunktionen abgeleitet: research problem, empirical object, method, theoretical, application. Da auch die Entwicklung und Dokumentation einer Vorgehensweise zur Berücksichtigung von Zitationsfunktionen bei der Analyse thematischer Strukturen einen Schwerpunkt der vorliegenden Arbeit darstellt, werden konzeptionelle Überlegungen zur Kodierung, das verwendete Kategoriensystem, die Vorgehensweise der Kodierung und der

Erstellung der Netzwerkkarten in Abschnitt 3.2 und 3.3 detailliert beschrieben.

Um einen empirischen Einblick in die Funktionalität von Zitierungen zu erhalten, werden die für Phase eins ausgewählten neun Publikationen als Volltexte in das Datenanalyseprogramm MaxQDA geladen. MaxQDA ermöglicht eine komfortable, übersichtliche und nachvollziehbare Kodierung der in den Artikeln enthaltenen Referenzen. Neben der epistemischen Funktion werden noch zwei weitere Dimensionen kodiert: der Ort des kodierten Segments innerhalb der Publikation nach Kapiteln und ob es sich um eine Selbstzitation handelt oder nicht. Um die Richtigkeit und Konsistenz der Kodierung zu gewährleisten werden alle Dokumente in einem iterativen Verfahren mehrfach durchgesehen und alle kodierten Segmente mehrfach geprüft. Die so erhobenen Daten können anschließend exportiert und mittels (deskriptiver) statistischer Verfahren analysiert werden. Die Ergebnisse geben einen Einblick in die Verteilungen von Zitationen nach verschiedenen Funktionen und ermöglichen Aussagen über die Abgrenzung feldinterner Forschungsansätze. Die Ergebnisse der Analyse werden in Abschnitt 4.1 erläutert.

Um die Implikationen zu untersuchen, die unterschiedliche Zitationsfunktionen auf das bibliometrische Mapping thematischer Strukturen haben könnten, werden in Phase zwei Netzwerkkarten erstellt und ausgewertet. Grundlage der Netzwerkkarten bildet ein Zeitschriftenartikel, der als Seed Dokument den Kern der Netzwerke bildet. Ausgehend von diesem Seed Dokument werden in Web of Science alle den Artikel zitierenden und alle in dem Artikel zitierten Quellen ermittelt. Die Referenzen aller so ermittelten Publikationen werden entsprechend ihrer epistemischen Funktion in MaxQDA kodiert. Nach entsprechender Aufbereitung der Daten können aus den kodierten Referenzen Netzwerkkarten für jede der epistemischen Funktionen erstellt werden. Das gibt einen ersten Einblick in die Verteilung und Strukturierung der Zitationsbeziehungen unterschiedlicher epistemischer Funktionen. Um Aussagen über die thematische Mehrdimensionalität des untersuchten Feldes treffen zu können, werden für alle Netzwerke Topics aus den Inhalten der Volltexte erhoben. Diese werden anschließend interpretiert und vergleichend analysiert. Die Darstellung der Ergebnisse aus der Netzwerkanalyse erfolgt in Abschnitt 4.2.

Das letzte Kapitel dient der abschließenden Reflektion der Vorgehensweise und der Zusammenfassung der Ergebnisse (Kapitel 5). Außerdem werden einige Punkte für an die behandelte Thematik anschließende Forschungsarbeiten in Aussicht gestellt.

2. Theoretische Einordnung und Stand der Forschung

Die vorliegende Arbeit beschäftigt sich mit (1) der Analyse epistemischer Funktionen von Zitationsbeziehungen und (2) deren Auswirkungen auf das bibliometrische Mapping thematischer Strukturen. Für die Einordnung in einen umfassenderen Forschungskontext sind daher sowohl die bestehende Literatur zu Zitationsfunktionen und deren Klassifikation als auch bisher erschienene Arbeiten zum Identifizieren und Abgrenzen von Forschungsfeldern mithilfe bibliometrischer Daten relevant. Darüber hinaus schließt die Arbeit an Erkenntnisse der (sozialwissenschaftlichen) Wissenschaftsforschung an.

Im Folgenden werden daher in drei Unterkapiteln die relevanten Stränge des Forschungskontexts dargestellt. In 2.1 wird ein kurzer Abriss über die einflussreichsten theoretischen Einsichten der Wissenschaftsforschung gegeben um die Arbeit in einen theoretischen Rahmen einzubetten. In 2.2 wird der Forschungsstand zum bibliometrischen Mapping thematischer Strukturen zusammengefasst und insbesondere die verschiedenen methodischen Vorgehensweisen dargestellt. In 2.3 werden einige frühere Arbeiten zu den verschiedenen Funktionen von Zitierungen dargestellt. Im Mittelpunkt stehen dabei vor allem die verschiedenen methodischen Vorgehensweisen und die Ausarbeitung unterschiedlicher Kategoriensysteme zur Klassifikation von Zitationsfunktionen.

2.1. Theoretische Einordnung

Wissenschaftsforschung, als Meta-Wissenschaft, also Wissenschaft über Wissenschaft, beschäftigt sich mit den Bedingungen und Praktiken innerhalb der Produktion, Kommunikation und Rezeption wissenschaftlicher Erkenntnis. Als interdisziplinäres Forschungsfeld unterliegt sie Einflüssen aus Philosophie, Anthropologie, Geschichtswissenschaft und nicht zuletzt der Soziologie. Eine der ersten soziologischen Untersuchungen zur Produktion wissenschaftlichen Wissens geht auf Ludwik Fleck (2012 [1935]) zurück. In *Entstehung und Entwicklung einer wissenschaftlichen Tatsache: Einführung in die Lehre vom Denkstil und Denkkollektiv* hebt Fleck die zentrale Rolle von sozialen Faktoren bei der Produktion wissenschaftlichen Wissens hervor. Wissenschaftliches Arbeiten organisiert sich in „Denkkollektiven“, die durch jeweils spezifische „Denkstile“ zusammengehalten werden. Ein in einem Denkkollektiv geteilter Denkstil umfasst nach Fleck die nicht artikulierten Überzeugungen der Wissenschaftler.

Viele Ideen und Ansätze von Fleck hat Thomas S. Kuhn (2014 [1962]) in *The Structure of Scientific Revolution* aufgegriffen oder weiterentwickelt. Bei Kuhn stehen die Prozesse, in denen wissenschaftliche Erkenntnisse produziert werden im Vordergrund. Anders als Fleck sieht Kuhn wissenschaftlichen Fortschritt nicht als kontinuierliche Weiterentwicklung, sondern als etwas, das sich in radikalen Brüchen, in „wissenschaftlichen Revolutionen“, abspielt. Innerhalb von wissenschaftlichen Revolutionen werden bestehende Erklärungsmodelle und in der Fachwelt anerkannte konkrete Problemlösungen durch neue ersetzt: Es findet ein Paradigmenwechsel statt.

Robert K. Merton hat sich im Rahmen seiner wissenschaftssoziologischen Arbeiten unter anderem mit den normativen Strukturen von Wissenschaft auseinandergesetzt. Seine Ausführungen zum „Ethos der Wissenschaften“ (Merton, 1942) war besonders einflussreich. Der Ethos beschreibt die institutionell verankerten und von Wissenschaftlern als wissenschaftliches Gewissen verinnerlichten Werte und Normen. Neben den inhaltlichen Merkmalen des Ethos galt Mertons Interesse vor allem auch der Rolle von wissenschaftsspezifischen Werten und Normen für das Fortschreiten der Wissenschaft und der verschiedenen ausgeprägten Verinnerlichung des Ethos bei Wissenschaftlern. Während Mertons institutionalistischem/normativem Ansatz in den 1950er bis 1970er Jahren noch viel Aufmerksamkeit zuteil wurde, verschob sich der Diskurs zum Ende dieser Zeit in eine andere Richtung. Erkenntnistheoretische Fragestellungen im Rahmen einer sozial-konstruktivistischen Perspektive übernahmen die Bühne der Wissenschaftsforschung (vgl. Panofsky, 2010).

Besonders die sogenannten Laborstudien erfuhren ab dieser Zeit zunehmend an Aufmerksamkeit. Darunter die Arbeiten von Bruno Latour und Steve Woolgar, die mit *Laboratory Life: The Construction of Scientific Facts* 1979 die erste Studie dieser Art vorbrachten. Im Rahmen der ethnografisch angelegten Studie wurden die Wissenschaftler des kalifornischen Salk Instituts direkt am Ort ihrer wissenschaftlichen Arbeit beobachtet: in ihrem Labor. Latour und Woolgar wollten so herausfinden wie wissenschaftliches Wissen in der alltäglichen Praxis hergestellt wird. Dabei schwingt die Annahme mit, dass zwischen alltäglicher wissenschaftlicher Praxis und denen ihr zugrunde liegenden methodischen Regeln eine Kluft besteht (vgl. Lengersdorf und Wieser, 2014, S. 100f).

Auch Karin Knorr Cetina veröffentlicht mit *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science* kurz darauf (1981) eine Laborstudie, die ähnliche Fragen aufwirft. Ebenfalls im Rahmen ethnografischer Erkundungen untersucht Knorr Cetina die Praktiken der Wissenschaft an den Orten ihrer Entstehung. Die Konfiguration von Objekten, die Konstruktion des epistemischen Subjekts und die Rolle des Labors sind dabei die zentralen Elemente ihrer Beobachtungen. In ihrem Nachfolgewerk *Epistemic Cultures: How the Sciences make Knowledge* (1999) postuliert sie anhand der vergleichenden Analyse von Molekularbiologie und Hochenergiephysik die These der Verschiedenheit unterschiedlicher Wissenskulturen. Die unterschiedlichen Wissenskulturen manifestieren sich je in Form von feldspezifischen Prozessen der Wissensgenerierung, die auf verfestigten Strategien, Orientierungen und Praktiken aufbauen (vgl. Kirschner, 2014, S. 131).

Jochen Gläser formuliert in seiner Habilitationsschrift das Konzept der „wissenschaftlichen Produktionsgemeinschaften“ aus (Gläser, 2006). Damit stellt er ein integratives Modell vor, das sowohl Aspekte der institutionalistischen als auch der sozial-konstruktivistischen Wissenschaftsforschung einfängt, um vor allem den Wechselwirkungen von Mikro- und Makroprozessen Rechnung zu tragen. Nachdem Gläser zunächst erläutert warum Konzepte wie „Markt“, „Organisation“ oder „Netzwerk“ für die Beschreibung der Wechselwirkungen innerhalb der Produktion wissenschaftlichen Wissens nicht funktionieren, führt er in das von Renate Mayntz und Fritz Scharpf (1995) vorgeschlagene Programm eines „akteurszentrierten Institutionalismus“ ein. Bei Mayntz und Scharpf findet Gläser auch das Konzept „kollektiver Produktionssysteme“, auf dem er im Weiteren aufbaut. Nach Gläser ist die Produktion wissenschaftlichen Wissens eine besondere Form eines solchen kollektiven Produktionssystems, nämlich eine *Produktionsgemeinschaft*. „Die Produktion beruht auf autonomen Entscheidungen der Produzenten darüber, welcher Beitrag gebraucht wird, dass sie diesen Beitrag erzeugen können, und wie sie dabei vorgehen. Diese Entscheidungen orientieren sich am gemeinsamen Wissensbestand der Gemeinschaft, der zugleich Arbeitsgegenstand, Arbeitsmittel und kollektives Produkt ist“ (ebd., S. 263). Später führt Gläser den Begriff der „epistemic regime’s“ (Gläser u. a., 2018) ein, die den gemeinsamen Wissensbestand, die Praktiken und die Struktur wissenschaftlicher Produktionsgemeinschaften zusammenführt. Epistemische Regimes sind nach Gläser die stabilen Arrangements von Praktiken

der Wissensproduktion und die sozialen Strukturen, in denen diese Praktiken durchgeführt werden (Gläser u. a., 2018, S. 3). Das eingeführte Konzept epistemischer Regime soll dabei helfen verschiedene wissenschaftliche Produktionsgemeinschaften vergleichend zu analysieren.

2.2. Forschungsstand thematisches Mapping

Vergleichende Analysen lassen sich in der Wissenschaftsforschung, bis auf wenige Ausnahmen, bis dahin aber kaum finden. Meist bleibt es bei der Analyse einzelner wissenschaftlicher Felder oder Disziplinen und selten werden Ergebnisse systematisch zusammengeführt. Neben der sozialwissenschaftlichen Wissenschaftsforschung hat sich insbesondere die Bibliometrie als Forschungsfeld etabliert, das sich mit der Identifikation und Abgrenzung thematischer Strukturen auseinandersetzt und jüngst einen besonderen Fokus auch auf vergleichende Analysen legt (vgl. hierzu insbesondere die Diskussion in *Scientometrics* 111(2) und 111(3)). Innerhalb der Bibliometrie haben sich ganz eigene Methoden entwickelt wissenschaftliche Felder voneinander abzugrenzen und thematische Strukturen zu identifizieren (vgl. Gläser, Glänzel und Scharnhorst, 2017). Häufig werden dazu die in bibliographischen Daten vorhandenen Informationen benutzt um thematisch naheliegende Publikationen zu identifizieren und nach Forschungsthemen zu clustern¹. Seit den ersten Versuchen der Identifikation und Abgrenzung thematischer Strukturen anhand bibliometrischer Daten durch Henry Small und seine Kollegen (Small, 1973; Small, 1977; Griffith u. a., 1974; Small und Griffith, 1974) hat sich an der grundlegenden Vorgehensweise nicht viel geändert: Eine Menge an Publikationen wird ausgewählt und danach mittels algorithmischer Verfahren analysiert, wobei verschiedene in der Publikation enthaltene Eigenschaften zugrunde gelegt werden können. Für die Analysen kommen meist diejenigen Publikationen in Frage, die in großen Publikationsdatenbanken wie dem Web of Science oder Scopus indexiert sind. Oftmals wird nicht der Inhalt der gesamten Datenbank für die Analyse ausgewählt, sondern es werden weitere, inhaltlich begründete, Einschränkungen vorgenommen. So ist es üblich die Auswahl nach bestimmten Suchabfragen einzuschränken. Je nach Fragestellung und Fokus des jeweils vorliegenden Projekts werden dann beispielsweise nur Publikationen ins Sample aufgenommen die bestimmte Wörter in Titel, Schlagwörtern oder Abstracts enthalten oder nur Publikationen aus bestimmten Fachzeitschriften ausgewählt. Aus den Datenbanken lassen sich dann die Metadaten der ausgewählten Publikationen exportieren, häufig auch mit Zitationsangaben. Nachdem die Daten ausgewählt und exportiert sind, lassen sich verschiedene Methoden auf sie anwenden.

Small und Kollegen schlugen bereits in den 1970er Jahren vor, Kozitationen (co-citation) für die Identifizierung und Abgrenzung von Forschungsfeldern zu nutzen (Small, 1973; Griffith u. a., 1974; Small und Griffith, 1974). Eine Kozitation stellt eine Verbindung zwischen zwei Aufsätzen her, die gemeinsam in später erschienenen Aufsätzen zitiert werden. Dahinter steht die Annahme, dass zwei Aufsätze, je häufiger sie gemeinsam zitiert werden, desto enger inhaltlich beieinander liegen. Um Kozitationshäufigkeiten von zwei Aufsätzen zu ermitteln wird der Pool der sie jeweils zitierenden Literatur daraufhin ausgewertet, ob die beiden Aufsätze häufig gemeinsam zitiert werden.

Nach der selben Logik funktioniert auch die Auswertung von Autoren-Kozitation (author co-citation), die 1981 von Howard D. White und Berverly C. Griffith eingeführt wurde. Anstatt die Kozitation zweier bestimmter Aufsätze zu analysieren, dient hier das gemeinsame Zitieren von Autoren als Grundlage

¹Einleitend in ein Special Issue über vergleichende Ansätze bei der Analyse thematischer Strukturen beschreiben Jochen Gläser, Wolfgang Glänzel und Andrea Scharnhorst (2017) die grundlegende Vorgehensweise solcher Arbeiten und fassen die wichtigsten Ansätze und Arbeiten kurz zusammen. Im Folgenden beziehe ich mich – unter Angabe der Primärquellen – auf den Überblick, den die Autoren dort liefern. Die Ausführungen der Autoren sind auf S. 982 - 986 nachzulesen.

der Analyse. Dadurch können Rückschlüsse darauf gezogen werden, welche Autoren sich inhaltlich und/oder methodisch nahestehen. White und Griffith bezeichnen das Kartieren bibliographischer Informationen mittels Autoren-Kozitation als Rekonstruktion des „consensus on past literature“ (White und McCain, 1998, S. 328). In Zitationen manifestiert sich über Zeit die kollektive Wahrnehmung der zitierten Literatur durch die sie zitierenden Autoren. Und diese kollektive Wahrnehmung drückt sich dann eben auch in den Clusterdiagrammen aus, die sich aus der Analyse ergeben. Die Nähe der Autoren kann sich dabei durch die Ähnlichkeit der bearbeiteten Themen oder der gepflegten Arbeitsweise ergeben (vgl. ebd.).

Wie bei der Kozitationsanalyse das gemeinsame Auftreten von Artikeln in Referenzlisten gemessen wird und bei der Autoren-Kozitationsanalyse das gemeinsame Auftreten von Autoren in Referenzlisten gemessen wird, wird bei der Kowortanalyse (co-word analysis) das gemeinsame Auftreten von Wörtern in einem Dokument für die Analyse der thematischen Strukturen genutzt. Aufbauend auf den methodologischen Überlegungen der Akteur-Netzwerk-Theorie entwirft Michel Callon (1983) die Kowortanalyse als Alternative zu zitationsbasierten Analyseverfahren. Vor diesem Hintergrund liegt der Fokus der Analyse auf Translationsprozessen anhand derer Probleme und vor allem die Beziehungen zwischen Problemen – also Problemnetzwerke („problematic networks“, Callon u. a., 1983, S. 194) – identifiziert werden können. Im Gegensatz zu zitationsbasierten Analysen ist die Kowortanalyse auf alle Textsorten anwendbar. Angewandt wurde die Kowortanalyse zum Beispiel zur Beschreibung der internen intellektuellen Strukturierung innerhalb der Biotechnologie (vgl. Rip und Courtial, 1984), in der Energieforschung (vgl. Tijssen, 1992) und für neuronale Netze (vgl. van Raan und Tijssen, 1993).

Eine weitere Grundlage zur Identifizierung thematischer Strukturen in der Wissenschaft können die Zitate liefern, die zwischen Zeitschriften ausgetauscht werden (journal-to-journal citation). Wissenschaftliche Fachzeitschriften sind mit der Ausdifferenzierung der Wissenschaften auch selbst immer ausdifferenzierter geworden. Neben einigen interdisziplinären Fachzeitschriften gibt es viele, die sich sehr spezialisierten Gebieten widmen. Es ist daher für die Kartierung der Wissenschaftslandschaft sehr aufschlussreich, in welchen Zeitschriften die Artikel welcher Zeitschriften zitiert werden, bzw. von welchen Zeitschriften Zitate auf die Artikel welcher Zeitschriften ausgehen. Loet Leydesdorff (1986) schlägt die Analyse von Zeitschriftenzitationen für die Beschreibung der kognitiven Differenzierung der Wissenschaften vor, die als Indikator für das dienen kann, was Leydesdorff als qualitative Veränderung bezeichnet. Einen besonderen Vorteil sieht Leydesdorff darin, dass Zeitschriftenzitationen in aggregierter Form vorlägen und deshalb kaum durch die Absichten der Autoren oder von politischen Entscheidungsträgern beeinflusst werden könnten (vgl. Leydesdorff, 1986, S. 106). Die Beziehung von Zitierungen zwischen Zeitschriften wurde sowohl auf der Ebene des gesamten Web of Science durchgeführt, als auch für einzelne Disziplinen (vgl. Leydesdorff und Cozzens, 1993; Leydesdorff, 2004; Leydesdorff und Rafols, 2009).

Die bibliografische Kopplung (bibliographic coupling) wird häufig als Gegenstück zur Kozitation beschrieben. Zwei Artikel sind bibliografisch gekoppelt, wenn beide mindestens einmal auf dieselbe dritte Arbeit referieren. Nachdem Maxwell Mirton Kessler die bibliografische Kopplung 1963 vorschlug und Samuel Schiminovich sie 1971 das erste Mal für die Identifizierung thematischer Strukturen nutzte, entfachte sich kein, etwa mit der Verwendung von Kozitationen vergleichbares, Interesse an der Nutzung der bibliografischen Kopplung zur Kartierung thematischer Felder (vgl. Gläser, Glänzel und Scharnhorst, 2017). Erst in den 1990er Jahren haben Wolfgang Glänzel und Hans-Jürgen Czerwon die Analyse bibliografischer Kopplung zur Identifikation angesagter Forschungsthemen wiederaufgenommen. Durch die Bestimmung angemessener Schwellenwerte konnten die Autoren über die bibliografische Kopplung die zentralen Dokumente („core documents“, Glänzel und Czerwon, 1995; Glänzel

und Czerwon, 1996) ausfindig machen. Auch Bo Jarvening (2001) nutzte bibliografische Kopplung um Kernkonzepte in der Herz-Kreislaufforschung zu identifizieren.

Aufbauend auf bisherigen Erfahrungen aus sowohl bibliometrischen als auch textanalytischen Verfahren entstehen Mitte der 2000er Jahre hybride Ansätze (vgl. Glenisson, Glänzel und Persson, 2005; Glenisson u. a., 2005; Janssens, Glänzel und Moor, 2007). Die Kombination von bibliometrischen und textanalytischen Analysemethoden wurde vor allem zur Verbesserung der Genauigkeit der einzelnen Methoden für die Identifizierung thematischer Strukturen entwickelt. So können zitationsbasierte Verfahren dabei helfen, die auf der Basis einer Kowortanalyse erhobenen Cluster zu spezifizieren.

Die meisten der bisher angeführten Arbeiten versuchten innerhalb eines nach bestimmten Kriterien eingegrenzten Sub-Sets an Publikationen thematische Strukturen zu ermitteln. Oft wurde sich dabei auf ein bestimmtes Fachgebiet, eine Disziplin oder einige bestimmte Fachzeitschriften beschränkt. Durch die fortgeschrittene Rechenleistung moderner Computer, die Entwicklung effizienterer Algorithmen und der Möglichkeit mancher Wissenschaftler auf lokale Kopien großer Zitationsdatenbanken zugreifen zu können, konnten seither auch Analysen erstellt werden, die sich auf die Wissenschaft als Ganzes beziehen (vgl. Boyack, Klavans und Börner, 2005; Klavans und Boyack, 2011; Waltman und van Eck, 2012).

Neben den recht klassisch bibliometrischen Verfahren, die bisher beschrieben wurden, etablieren sich zunehmend auch probabilistische Modelle für die Identifikation thematischer Strukturen innerhalb der Wissenschaftsforschung (z.B. Blei, 2012a; Blei, 2012b; Yau u. a., 2014). Für die Modelle werden Volltexte analysiert, da davon ausgegangen wird, dass die in Dokumenten besprochenen Themen anhand dem Vorkommen von Wörtern identifiziert werden können. Dabei wird angenommen, dass bestimmte Wörter eine relativ höhere Wahrscheinlichkeit haben im Dokument aufzutauchen, wenn es um ein bestimmtes Thema geht. Eine häufig verwendete Modell ist die „Latent Dirichlet Allocation“, die als relativ einfaches Verfahren unter den probabilistischen Topic Modelling-Ansätzen gilt und auf David M. Blei (2003) zurückgeht. Basierend auf der Vorgehensweise des machine learning werden Terme im Rahmen eines feststehenden Wortschatzes Themen zugeordnet, sodass der Algorithmus anhand der in den Dokumenten jeweils vorkommenden Wörter mit einer bestimmten Wahrscheinlichkeit die 'verborgenen' Themen des Dokuments ermitteln kann. Hauptkritikpunkt an dem Verfahren ist, dass die Anzahl der in den untersuchten Dokumenten enthaltenen Themen vorab festgelegt werden muss (vgl. Gläser, Glänzel und Scharnhorst, 2017).

Das Identifizieren thematischer Strukturen wissenschaftlichen Publikationsoutputs insgesamt oder innerhalb einzelner wissenschaftlicher Fachgebiete und Disziplinen ist für verschiedene Zielgruppen und für unterschiedliche Verwendungszwecke interessant. Zum einen sind die Ergebnisse für Wissenschaftsforscher relevant, die gerne mehr über die thematischen Strukturen erfahren möchten. Sie können damit die Ausdifferenzierung der Wissenschaften nachzeichnen, sie können das Entstehen neuer Forschungsfelder erkennen und beschreiben und sie können zu einem tieferen Verständnis von „Denkstilen“ und „Denkkollektiven“ (Fleck, 2012 [1935]), von „wissenschaftlichen Revolutionen“ und „Paradigmenwechseln“ (Kuhn, 2014 [1962]), von „Wissenskulturen“ (Knorr Cetina, 2002 [1999]) oder „epistemischen Regimes“ (Gläser u. a., 2018) innerhalb „wissenschaftlicher Produktionsgemeinschaften“ (Gläser, 2006) beitragen. Zum anderen ist die Identifizierung und Abgrenzung thematischer Strukturen wesentlich für bibliometrische Evaluationen. Da Zitationszahlen in verschiedenen Fachbereichen aufgrund unterschiedlicher Zitationskulturen und unterschiedlicher Größe der Fachgemeinschaften ganz unterschiedlich ausfallen, ist es inzwischen üblich durch Normalisierung Vergleichbarkeit herzustellen. Für die Normalisierung bibliometrischer Indikatoren ist es demnach wichtig die Grenzen wissenschaftlicher Fachgebiete definieren zu können, wobei die oben genannten Verfahren einen Beitrag leisten können. Die Verwendung der verschiedenen Verfahren zur Identifikation und Abgrenzung

von Forschungsthemen ist also sowohl für Wissenschaftsforscher interessant, als auch für diejenigen, die die bibliometrischen Daten und Indikatoren herstellen (Plattformanbieter und Bibliometriker), die die Indikatoren zu evaluativen Zwecken anwenden (Wissenschaftsmanager und Lenkungsgruppen) und die mittels der Indikatoren beurteilt werden (Wissenschaftler, Forschungsgruppen, etc., vgl. z.B. Leydesdorff, Wouters und Bornmann, 2016).

2.3. Forschungsstand Zitationsfunktionen

Obwohl die meisten der oben beschriebenen Verfahren zur Identifizierung und Abgrenzung thematischer Strukturen in der Wissenschaft zitationsbasiert sind, wird die Bedeutung einer Zitation als solche kaum reflektiert. Unterbelichtet bleiben deshalb die verschiedenen Funktionen, die Zitationen in wissenschaftlichen Arbeiten erfüllen können. Tatsächlich werden Zitationen aber in ganz unterschiedlichen Kontexten und Zusammenhängen genutzt und verweisen auf eine Vielzahl verschiedener thematischer Bezüge. Beispielsweise kann eine Zitation auf die Verwendung einer anderswo publizierten Methode verweisen, auf Informationen über den Forschungsgegenstand, auf Vorarbeiten, die die vorliegende Arbeit motivieren oder auf theoretische Einordnungen. Da auch in den bibliografischen Datenbanken keinerlei Informationen zu den verschiedenen Zitationsfunktionen enthalten sind, bleiben diese bei der Identifizierung thematischer Strukturen unberücksichtigt, bzw. vermischen sich in den genutzten Modellen.

Dass Zitationen in wissenschaftlichen Arbeiten unterschiedliche Zwecke erfüllen, bzw. auf verschiedene inhaltliche Aspekte bezogen sind, ist offensichtlich und auch erforscht (siehe Tahamtan und Bornmann, 2019 und Bornmann und Daniel, 2008 für eine ausführliche Zusammenfassung des Forschungsstands zu Zitationsfunktionen). Erste konzeptionelle Arbeiten zu den verschiedenen Funktionen, die Zitierungen im Rahmen wissenschaftlicher Aufsätze erfüllen können, erscheinen in den 1960er Jahren. Sowohl Eugene Garfield (1962) als auch Ben-Amin Lipetz (1965) entwerfen bereits eigene Systeme zur Klassifizierung verschiedener Zitationsfunktionen. Im Groben lassen sich für Arbeiten zu Zitationsfunktionen drei methodische Ansätze differenzieren: (1) Inhaltsanalysen, (2) Kontextanalysen und (3) Umfragen und Interviews. Besonders bezogen auf inhalts- und kontextanalytische Verfahren spielt neuerdings Automatisierung eine wichtige Rolle. Während inhalts- und kontextanalytische Verfahren eher versuchen die Beziehung zwischen zitierendem und zitiertem Dokument anhand der in den Dokumenten selbst enthaltenen Informationen zu analysieren, fokussieren sich Studien mit Umfrage- oder Interview-Design eher auf die Motive, die die Wissenschaftler selbst für ihr Zitationsverhalten angeben. Inhaltsanalysen richten sich dabei eher auf den Inhalt der zitierten Publikation und Kontextanalysen eher auf die Inhalte rund um das Zitat im zitierenden Dokument (vgl. ebd.). Im Folgenden sollen kurz einige wegweisende Studien für die drei jeweils verschiedenen methodischen Ansätze vorgestellt werden. Dabei werden nicht die Ergebnisse der Studien besonders in der Vordergrund gestellt, sondern vielmehr die Klassifikationen und Systematisierungen der erarbeiteten Zitationsfunktionen.

Bei Kontextanalysen ist es der Kontext in dem das Zitat steht, der im Fokus der Analyse steht. Durch die Analyse der Worte und ihrer semantischen Inhalte, die im zitierenden Dokument das Zitat einleiten und einbetten, soll herausgefunden werden wie die Beziehung zwischen zitiertem und zitierendem Dokument ist (vgl. z.B. Cronin, 1984, McCain und Turner, 1989). Kontextanalysen sind deshalb immer dokumenten-basiert und erfordern es (sieht man zunächst einmal von automatisierten/computergestützten Verfahren ab), dass das zitierende Dokument komplett gelesen und jede Stelle intellektuell geprüft wird.

Moravcsik und Murugesan liefern erste umfassende Ergebnisse aus einer Zitationskontextanalyse (Moravcsik und Murugesan, 1975; Murugesan und Moravcsik, 1978). Die Autoren erarbeiten fünf Kategorienpaare, die die Fragen beantworten, ob eine Zitation (1) „conceptual or operational“, (2) „organic or perfunctory“, (3) „evolutionary or juxtapositional“, (4) „confirmative or negational“, (5) „valuable or redundant“ (Moravcsik und Murugesan, 1975, S. 88) ist. Während die erste und dritte Dimension einen Einblick darin, gibt welcher Art die Beziehung zwischen zitierendem und zitiertem Dokument ist, sind die anderen Dimensionen eher auf die Qualität der zitierten Quelle gerichtet.

Chubin und Moitra (1975) entwerfen eine Klassifikation, die zwischen affirmativen und negierenden Zitationen unterscheidet, wobei die affirmativen Zitationen sich in vier Typen differenzieren und die negierenden zwei Unterklassen aufweisen. Innerhalb der vier zustimmenden Kategorien wird differenziert wie zentral, relevant, oder wichtig die zitierte Quelle für die vorliegende Arbeit ist. Innerhalb der zwei negierenden Kategorien wird unterschieden wie stark die Autoren von der in der Quelle erwähnten Position, Vorgehensweise oder Schlussfolgerung abweichen. Es geht darum zu unterscheiden, ob der Zitierende nur teilweise nicht mit der zitierten Quelle übereinstimmt und z.B. eine Korrektur der Vorgehensweise anbietet, oder ob er die Quelle für komplett falsch hält und eine ganz eigene Interpretation oder Lösung liefert.

Spiegel-Rösing (1977) hat dreizehn verschiedene Funktionen von Zitierungen klassifiziert. Darunter sind Kategorien wie: Verweise auf verwendete Methoden, Verweise auf Konzepte und Definitionen, Verweise auf verwendete Daten, Verweise auf vorausgegangene Arbeiten zur selben Forschungsfrage. Kategorien wie diese finden sich in vielen Arbeiten zum Thema, häufig unterschiedlich formuliert und mit verschiedenen Definitionen und Erklärungen, sowie in verschieden detaillierter Form.

Bei Krampen und Montada (2002) finden sich verschiedene Kategorien, die in anderen Klassifikationssystemen so nicht vorkommen. Darunter direkte Referenz auf konkrete empirische Ergebnisse, einfache Verweise ohne konkretere Angabe („compare here also“, „see also“, see, for example“, S. 69), Verweise auf Theorien oder Konzepte, Verweise auf Methoden oder z.B. die Verwendung bestimmter Datensätze. Es wurde ebenfalls erfasst, ob es sich um ein direktes wörtliches Zitat handelt oder Abbildungen oder Tabellen direkt aus der Quelle übernommen wurden. Mit elf verschiedenen Kategorien ist das System durchschnittlich ausdifferenziert.

Die in diesen Studien thematisierten Aspekte geben einen Einblick in die Dimensionen auf deren Codierung die einzelnen Wissenschaftler und Arbeitsgruppen Wert gelegt haben. Insgesamt fällt auf, dass in den meisten Fällen eigene Kategoriensysteme entworfen wurden und nur selten die Klassifikationen von anderen nachgenutzt wurden. Nur selten wurde versucht Ergebnisse zu replizieren. In vielen Kategoriensystemen gibt es zwar Überschneidungen einzelner Elemente, dennoch wird in fast jeder Arbeit nochmal ein besonderes Schlaglicht auf einen bestimmten Aspekt geworfen. Gemeinsam ist fast allen Studien, dass wie auch immer geartete positive und negative Zitationen unterschieden werden. Außerdem scheint es vielen Autoren wichtig zu sein zwischen beiläufigen, recht unspezifischen Zitationen und konkreten Zitationen zu unterscheiden. Die Klassifizierung der konkreten inhaltlichen Objekte, auf die sich die Zitationen beziehen, wie z.B. eine methodische Vorgehensweise, ein bestimmtes empirisches Objekt oder Bemerkungen zur generellen Beschreibung des Forschungsproblems finden sich ebenfalls in vielen der oben genannten Studien.

Inhaltsanalysen unterscheiden sich vom Fokus her von Kontextanalysen, da hier (ausschließlich) der semantische Inhalt der zitierten Passage oder des zitierten Dokuments bei der Charakterisierung der Zitationsfunktionen berücksichtigt wird. Einige inhaltsanalytische Studien untersuchen wie bestimmte, z.B. sehr einflussreiche oder sehr umstrittene Arbeiten rezipiert werden oder die Rezeption der Arbeiten bestimmter Autoren. So untersucht z.B. Cole (1975) Artikel, die Arbeiten des einflussreichen

(Wissenschafts-)Soziologen Robert K. Merton zitieren und bekommt dabei heraus, dass die meisten Rezeptionen ihn in einer zeremoniellen Art zitieren.

Garfield (1978), Garfield und Welljams-Dorof (1990) und Kochan und Budd (1992) untersuchen die Zitationen, die auf sehr umstrittene Arbeiten fallen. Bei den Studien ging es darum, den Einfluss betrügerischer Arbeiten zu messen und herauszufinden, warum die Arbeiten dennoch zitiert wurden und in welcher Weise. Wurden die Zitate positiv, negativ oder neutral im Text eingebettet und wie hoch war der Anteil von Selbstzitationen. Grundlage waren die Zitationen, die auf Arbeiten referierten, bei denen bekannt geworden ist, dass entweder die Daten und Ergebnisse gefälscht wurden oder sonstiges wissenschaftliches Fehlverhalten der Autoren bekannt geworden war.

Oppenheim und Renn (1978) untersuchten besonders hoch zitierte Arbeiten (z.B. bekannte Artikel von Albert Einstein) und schlagen dabei eine mit sieben Kategorien recht einfache Typologie zur Klassifizierung der Zitationen vor. Gegenüber differenzierteren Typologien, wie die Autoren betonen, habe sie den Vorteil einfacher anwendbar zu sein und auch einfacher verständlich, also intersubjektiv nachvollziehbar, zu sein. Die Zitationen werden demnach einer der sieben Kategorien zugeordnet: (1) „historical background“, (2) „description of other relevant work“, (3) „supplying information or data, other than for comparison“, (4) „supplying information or data for comparison“, (5) „use of theoretical equation“, (6) „use of methodology“, (7) „theory or method not applicable or riot the best one“ (Oppenheim und Renn, 1978, S. 226).

Hooten (1991) benutzt mehrere der bisher vorgestellten Klassifikationssysteme (Murugesan und Moravcsik, 1978; Peritz, 1983; Chubin und Moitra, 1975; Spiegel-Rösing, 1977), da sie jeweils auf unterschiedliche Aspekte von Zitationsfunktionen eingehen. Daraus ergab sich ein Kategoriensystem mit 35 verschiedenen Klassen. Das Kategoriensystem wurde im Rahmen einer vergleichenden Analyse von hochzitierten und selten-zitierten Artikeln eingesetzt. Die zitierenden Aufsätze der beiden Gruppen wurden dann hinsichtlich ihrer Zitationsstruktur anhand der Klassifikation analysiert.

Einen interessanten Ansatz wählten auch Maričić u. a. (1998), weil sie zum ersten Mal den Ort innerhalb des zitierenden Dokument mitkodiert haben. Sie haben erhoben, in welchem Teil der zitierenden Arbeit das Zitat vorkommt: Einleitung, Methodenteil, Ergebnisse oder Diskussion. Darüber hinaus haben die Autoren in sechs Abstufungen erhoben, wie wichtig die zitierte Quelle für die vorliegende Arbeit ist. Sie konnten daran z.B. zeigen, dass kursorische und oberflächliche Zitate eher in der Einleitung vorkommen, wobei zentrale und bedeutungsvolle Quellen eher in Methoden- und Ergebniskapiteln zitiert werden.

Insgesamt zeigt sich für inhaltsanalytische genauso wie für kontextanalytische Studien, dass sehr viele verschiedene Kategoriensysteme zur Klassifikation von Zitationsfunktionen entworfen wurden. Auch wenn im Rahmen von Inhaltsanalysen anders vorgegangen wird als bei Kontextanalysen, lassen sich die klassifizierten Dimensionen kaum unterscheiden. Viele der erhobenen Kategoriensysteme sind sowohl durch inhalts- wie auch durch kontextanalytische Verfahren zu ermitteln. Wie v.a. auch die Studie von Hooten (1991) zeigt, sind die Kategoriensysteme zwischen Kontext- und Inhaltsanalysen zum Teil sogar austauschbar. Lediglich die Ermittlung des Zitationsorts innerhalb des zitierenden Dokuments scheint inhaltsanalytischen Verfahren vorbehalten.

Neben den textbasierten Ansätzen zur Ermittlung von Zitationsfunktionen gibt es auch noch Studien, die mittels Umfragen und Interviews etwas über das Zitationsverhalten von Wissenschaftlern herausfinden möchten. Die Autoren wissenschaftlicher Arbeiten sollen dabei ihre Intention für die jeweilige Zitierung selbst erläutern, bzw. anhand eines Kategoriensystems zuordnen. In einigen Studien wurden die Wissenschaftler gebeten die Gründe für das Zitieren in eigenen Worten zu formulieren (vgl. z.B. Brooks, 1985; Brooks, 1986), in anderen wurden die Wissenschaftler gebeten ihre Zitationen anhand eines vordefinierten Schemas zu klassifizieren (vgl. z.B. Vinkler, 1987; Cano, 1989). So sehen es viele

Studien vor, den Autoren Checklisten mit möglichen Gründen für eine Zitation vorzulegen, woraus die Autoren dann für spezifische Quellen ihrer eigenen Aufsätze Zutreffendes auswählen können. In den Studien wurden teilweise Kategorien aufgegriffen, die auch schon in Inhalts- und Kontextanalysen gefunden wurden, zum Teil aber auch Kategorien entwickelt, bzw. gefunden, die noch neu sind. So wurde beispielsweise angegeben, dass ein beruflicher oder privater Vorteil durch die Zitierung erwartet wird oder dass die Zitierung getätigt wurde, weil der Autor sich in einer Art Abhängigkeitsverhältnis zum Autor der zitierten Quelle befindet (Vinkler, 1987). Es konnte außerdem herausgefunden werden, dass Autoren Quellen zitieren, die sie persönlich besonders beeinflusst haben, z.B. weil sie die Hauptquelle einer Idee waren und dass Zitationen aus sozialen Gründen getätigt werden, z.B. wenn erwartet wird, dass der Autor des zitierten Artikels im Review Prozess einflussreich sein könnte (Shadish u. a., 1995). Textbasierte Analyseverfahren lassen Kategorisierungen dieser Art natürlich nicht zu. Die Informationsgrundlage reicht zu solchen Schlussfolgerungen schlicht nicht aus. Aus diesem Grund tragen Interview- und Surveystudien in jedem Fall für ein tiefergehendes Verständnis von Zitationsfunktionen bei. Einige Kategorisierungen von Zitationsfunktionen überschneiden sich aber auch mit denen aus rein textbasierten Analysemethoden. Cano (1989) nutzt zum Beispiel die Klassifikation von Moravcsik und Murugesan (1975) für die Erhebung. Wie Maričić u. a. (1998) versucht auch Cano den Zusammenhang zwischen dem Ort der Zitation, also in welchem Kapitel auf eine Quelle referiert wird, und der Funktion der Zitation herzustellen. Snyder und Bonzi (1998) widmen sich in einem vergleichenden Ansatz den unterschiedlichen Gründen für das Zitieren eigener, bzw. fremder Werke.

Seit den ersten empirischen Untersuchungen zu den verschiedenen Funktionen von Zitationen in wissenschaftlichen Veröffentlichungen wie sie oben skizziert wurden, hat sich hinsichtlich der technischen Voraussetzungen und Möglichkeiten einiges verändert. Nicht nur liegen die Zitations- und Metadaten wissenschaftlicher Publikationen in enormen Mengen in großen Datenbanken zum Download bereit, auch sind die Volltexte der Publikationen inzwischen häufig in maschinenlesbarer Form vorhanden und zugänglich (vgl. Tahamtan und Bornmann, 2018). Technologische Neuerungen wie diese, kombiniert mit technischen Weiterentwicklungen von computergestützten Verfahren und Verfahren des machine learnings und data minings ermöglichen eine automatisierte, oder zumindest computergestützte, Ermittlung von Zitationsfunktionen. Wo im Rahmen von kontext- und inhaltsanalytischen Verfahren bisher Publikationen unter Inkaufnahme eines enormen Arbeits- und Zeitaufwands (vgl. McCain und Turner, 1989) von vorne bis hinten gelesen werden und die darin enthaltenen Zitationen interpretiert werden mussten, können nun maschinelle Verfahren eingesetzt werden.

Die in diesem Kapitel dargestellten theoretischen, empirischen und methodischen Einsichten dienen der Einbettung des Forschungsvorhabens in aktuelle wissenschaftliche Diskurse. Durch die Reflektion bisheriger Forschungsarbeiten konnte gezeigt werden, dass das Mapping thematischer Strukturen auf Grundlage bibliometrischer Daten zwar ein viel bearbeitetes Forschungsgebiet darstellt, die systematische Berücksichtigung von Zitationsfunktionen dabei aber bisher unberücksichtigt bleibt. Die vorliegende Arbeit leistet demnach nicht nur einen Beitrag zum besseren Verständnis von Zitationsfunktionen und deren Einfluss auf thematische Mappings, sondern trägt auch zur Weiterentwicklung von Mappingverfahren bei. Darüber hinaus wird ein Beitrag zum besseren und tiefergehenden Verständnis der epistemischen Praktiken im untersuchten Forschungsgebiet der Invasionsbiologie geleistet.

3. Methodisches Vorgehen

Im Folgenden wird das methodische Vorgehen erläutert. Da die vorliegende Arbeit in zwei Teile gegliedert ist, ist der Abschnitt 3.1 Sampling in zwei Unterkapitel aufgeteilt. Im ersten Unterkapitel (3.1.1) wird die Auswahl der Publikationen für die Klassifizierung der epistemischen Funktionen von Zitierungen erläutert und begründet. Im zweiten Unterkapitel (3.1.2) wird die Auswahl der Publikation für die Netzwerkanalyse erläutert und begründet. In den Abschnitten 3.2 Klassifikation und 3.3 Netzwerkmodellierung wird dann das analytische Vorgehen eingehender beschrieben.

3.1. Sampling

An dieser Stelle soll ein aktuell laufendes Forschungsprojekt, das für die vorliegende Arbeit einige Bedeutung hat und insbesondere für das Sampling ausschlaggebend war, Erwähnung finden. Am Deutschen Zentrum für Hochschul- und Wissenschaftsforschung beschäftigt sich eine Arbeitsgruppe innerhalb des Projekts „Fachspezifische Formen von Open Science“ unter der Leitung von Dr. Theresa Velden mit der Frage, welche feldspezifischen Faktoren die Weitergabe epistemischer Ressourcen, d.h. Forschungsdaten, Computercode und andere im Forschungsprozess entstehende Materialien, beeinflussen und wie diese Beeinflussung zur erklären ist. Die Analyse feldspezifischer Faktoren lässt sich dabei auf die theoretische Überlegung zurückführen, dass Wissenserzeugung innerhalb „wissenschaftlicher Produktionsgemeinschaften“ (Gläser, 2006) abläuft und diese hinsichtlich ihrer „epistemische[n] Regimes“ (Gläser u. a., 2018) vergleichend analysiert werden können. Im Forschungsprojekt werden innerhalb von vier abgrenzbaren Feldern (Invasionsbiologie, Solar Physics, Cluster Science, Synthetic Chemistry) Experteninterviews durchgeführt und Methoden der ethnographischen Feldforschung eingesetzt um Forschungspraktiken und Entscheidungen zur Weitergabe epistemischer Ressourcen zu erklären. Im Projekt werden auch Verfahren des bibliometrischen Mappings zur Erfassung der Variation epistemischer Bedingungen innerhalb der vier Felder verwendet. Das Projekt ist aus zwei Gründen für die vorliegende Arbeit besonders relevant: Innerhalb des Projekts liegen vertiefte Kenntnisse über die vier untersuchten Forschungsfelder vor, von denen die vorliegende Arbeit profitiert. Innerhalb des Projekts wurden Felddatensätze für die bibliometrischen Mappings der vier Felder erstellt, von denen einer in der vorliegenden Arbeit nach-genutzt wird.

Grundgesamtheit Felddatensatz

Für die Erhebung der Grundgesamtheit stütze ich mich auf den „Felddatensatz Invasionsbiologie“ der im Rahmen des Forschungsprojekts „Fachspezifische Formen von Open Science“ unter der Leitung von Dr. Theresa Velden am Deutschen Zentrum für Hochschul- und Wissenschaftsforschung (DZ-HW) durchgeführt wird. Die Datengrundlage für den Felddatensatz für die Invasionsbiologie ist durch Suche mittels einer lexikalischer Abfrage („lexical query“, siehe A.1, Dokumentation Lexical Query Felddatensatz Invasionsbiologie, S. 66) in der „advanced search“ Suchmaske der web-basierten Version des Web of Science am 26. Juni 2019 erzeugt worden, wobei das Zeitfenster 2000–2019 gewählt

wurde. Berücksichtigt wurden dabei nur Publikationen, die sich im Science Citation Index Expanded oder im Social Sciences Citation Index befanden. Im nächsten Schritt wurden die bibliographischen Metadaten sowie die Zitationsbeziehungen mittels SQL-Abfragen aus der Web of Science Datenbank des Kompetenzzentrums Bibliometrie heruntergeladen. Für den resultierenden Datensatz wurde das Zitationsnetzwerk erstellt und die größte zusammenhängende Komponente („giant component“) dieses Netzwerkes bildet den Felddatensatz. Durch Vernachlässigung kleinerer Netzwerkkomponenten und einzelner, unverbundener Publikationen erhöht sich die Präzision der Abgrenzung des Feldes, da durch die lexikalische Abfrage auch irrelevante, nicht der Invasionsbiologie zuzurechnende Publikationen erfasst werden. Von anfangs 65.046 Publikationen verbleiben schließlich 53.524 Publikationen, die den Felddatensatz ausmachen, wobei von diesen wiederum nur 49.541 einem thematischen Cluster zugeordnet werden konnten. Die beiden Clusterlösungen (Leiden7, Leiden15) ordnen die Publikationen des Felddatensatzes 7 bzw. 15 thematischen Clustern zu, wobei all jene Publikationen, die zwar in der giant component enthalten sind, sich aber algorithmisch nicht einem dieser Themencluster zuordnen ließen, einem weiteren „Nullcluster“ (das sind die verbleibenden 3.983 Publikationen mit Clusterindex = -1) zusammengefasst¹.

Die Leiterin der Arbeitsgruppe, Frau Dr. Theresa Velden, hat mir freundlicherweise die IDs der für den Felddatensatz mittels der Suchabfrage in Web of Science ermittelten Publikationen zur Verfügung gestellt. Über die WoS-IDs (UTs/Accession Numbers) konnte ich wiederum die Metadaten der Publikationen im Web of Science abfragen und für das weitere Sampling-Verfahren aufbereiten². Es konnten auf diese Weise 49.536³ Publikationen ermittelt werden (vgl. 3.1). Sie stellen die Grundgesamtheit für das Sampling und die Analysen dar.

	Anzahl Publikationen
Leiden1	19.095 Publikationen
Leiden2	9.771 Publikationen
Leiden3	5.839 Publikationen
Leiden4	5.459 Publikationen
Leiden5	3.710 Publikationen
Leiden6	3.303 Publikationen
Leiden7	2.359 Publikationen
insgesamt	49.536 Publikationen
davon 2018	4.396 Publikationen
davon 2009	2.503 Publikationen

Tabelle 3.1.: Übersicht Abfrage WoS-IDs

Von diesen 49.536 wurden die vollständigen Metadaten aller Publikationen aus den Jahren 2018 und 2009 in mehreren Schritten exportiert und in zwei Datensätzen zusammengeführt. Die 4.396 Publikationen mit Erscheinungsjahr 2018 dienen als Grundlage der Auswahl der Publikationen für die Klassifikation von Zitationsfunktionen (Phase 1 des empirischen Teils). Die 2.503 Publikationen mit Erscheinungsjahr 2009 dienen als Grundlage für die Auswahl der Publikation für die Netzwerkanalyse (Phase 2 des empirischen Teils)⁴.

¹Bei den Publikationen im Nullcluster lässt sich also nicht ohne Weiteres von einem thematischen Bezug zueinander ausgehen.

²Die gesamte Abfrage der Publikationen ist auf dem der Arbeit beiliegenden Datenträger dokumentiert. Die Datei ist unter dem Namen „suche_leiden7“ im nicht-proprietären txt-Format abgegelgt. Da die Dokumentation der Abfrage über 1.000 Seiten einnimmt, wurde von einer Aufnahme in den Anhang abgesehen.

³Die Differenz zu den 49.541 Publikationen, die ursprünglich in der giant component des Felddatensatzes enthalten sind, lassen sich auf vereinzelte Veränderungen in der Datenbank (Web of Science) zurückführen.

⁴Die Einschränkung der Publikationen mit Erscheinungsjahr 2018 und 2009 ist inhaltlich begründet. Weitere Erläuterungen dazu sind in den Unterkapiteln zum jeweiligen Sampling (3.1.1 und 3.1.2) dokumentiert.

3.1.1. Sampling Phase 1 – Klassifizierung von Zitationsfunktionen

Sampling-Ziel

Ziel des Sampling-Prozesses für die Klassifizierung von Zitationsfunktionen war es neun Publikationen für die Analyse der epistemischen Funktionen von Zitationen auszuwählen. Die Größe des Samples orientierte sich dabei an pragmatischen Überlegungen hinsichtlich des Arbeitsaufwands und der zur Verfügung stehenden Zeit. Neben forschungspraktischen Überlegungen standen aber auch theoretische Überlegungen zur Varianz von Forschungsansätzen im Mittelpunkt der mehrstufigen komplexen Auswahl.

Die Forschungsprozesse und -praktiken innerhalb ein und desselben wissenschaftlichen Spezialgebietes wie z.B. der Invasionsbiologie variieren z.T. stark. Es lassen sich innerhalb eines Gebietes, das durch eine übergreifende Fragestellung (Prozess der Invasion in der Ökologie) definiert ist, epistemische Teilgemeinschaften ausmachen, die sich voneinander in der Art und Weise, wie sie empirische Evidenz generieren oder auf alternative Weise, durch theoretische Überlegungen, wissenschaftliche Ergebnisse erzielen, unterscheiden. Die einen verfolgen stark empirische Ansätze, wobei sie Feldbeobachtungen machen oder in Laboren experimentieren, die nächsten setzen auf Modellrechnungen, die sie computer-gestützt anhand großer Datensätze durchführen und wieder andere Arbeiten sind sehr anwendungs-, bzw. policy-orientiert. Dies legt nahe, für die vergleichende Analyse der epistemischen Funktionen von Zitierungen, die feldinterne Heterogenität epistemischer Praktiken nicht zu ignorieren, sondern im Forschungsdesign zu berücksichtigen. Im Folgenden wird eine Auswahl der wichtigsten bzw. am deutlichsten abgrenzbaren, epistemischen Ansätze in der Invasionsbiologie, die für das Sampling herangezogen werden können, näher beschrieben.

Unter biologischer Invasion versteht die Invasionsbiologie die menschlich verursachte Verbreitung von Spezies in Gebieten, in denen sie ursprünglich nicht beheimatet waren. Die Invasionsbiologie erforscht die Invasion biologischer Spezies als Prozess – sie beschreibt ihn, klärt kausale Bedingungen und Mechanismen auf, trifft Vorhersagen, untersucht Auswirkungen und informiert Management und Politik, um die gezielte Beeinflussung von Invasionsprozessen zu unterstützen. Die Invasionsbiologie entwickelte sich als wissenschaftliches Spezialgebiet in der Ökologie seit Mitte bis Ende der 1980er Jahre. Mehrere wissenschaftliche Zeitschriften mit einem thematischen Fokus auf biologische Invasion wurden in den 1990er Jahren und dem frühen 21. Jahrhundert gegründet (vgl. Davis, 2006). Am Rand des Forschungsfeldes angesiedelt sind Forschungsarbeiten, die Invasion nicht explizit als Prozess adressieren, aber in gewissem Maße als der Invasion bewusst ('invasion aware') bezeichnet werden können. Das sind z.B. Studien, die die Biologie und oder die Ökologie invasiver Spezies allein oder zusammen mit einer nicht-invasiven Spezies untersuchen, ohne sie in dem unmittelbaren Kontext einer Invasionsdynamik zu betrachten.

Für die Invasionsbiologie lassen sich zumindest vier Forschungsansätze deutlich voneinander abgrenzen: empirische Studien auf Mikroebene, Modellierungsstudien auf Makroebene, theoretische Beiträge und anwendungsorientierte Arbeiten (siehe 3.2).

Die Einteilung in diese vier Typen von Forschungsansätzen ist zunächst einmal eine grobe Gliederung, der die Forschungsaktivitäten der Invasionsbiologie strukturierenden epistemischen Kontexte (vgl. ebd.). Es ist nicht auszuschließen, dass auch Mischformen der einzelnen Ansätze vorkommen, oder Typen, die hier noch keine Berücksichtigung finden. Der Forschungsansatz „empirisch“ ließe sich sicherlich noch weiter differenzieren und umfasst vermutlich einen Großteil der invasionsbiologischen Studien. Die hier vorgenommene Unterteilung soll zunächst vorrangig für die Strukturierung des Samples dienen. Da rein theoretische Arbeiten in der Invasionsbiologie nur selten vorkommen, wird dieser

Typ	Themen und Zielsetzung	Häufige Methoden
Empirisch (mikro)	Prozesse und Mechanismen der Invasion, Eigenschaften invasiver Arten, kausale Wirkzusammenhänge	Common Garden Experimente, Greenhouse Experimente, Genetische Analysen, Surveys und Feldbeobachtungen
Modellierung (makro)	Ausbreitung invasiver Spezies, Historie, Vorhersagen, Aufklärung von Zusammenhängen	Computersimulation, statistische Analyse großer Datensätze
Theoretisch (makro)	Literaturbasierte Metaanalysen, konzeptionelle Ansätze	Invasionshypothesen, Faktoren die die Invasivität beeinflussen
Anwendungsorientiert	Management invasiver Spezies, Minimalisierung von Risiken und Auswirkungen	Interviews, Umfragen, Bewertungskonzepte

Tabelle 3.2.: Übersicht und Kennzeichen der verschiedenen Forschungsansätze in der Invasionsbiologie

Typ vom Sample ausgeschlossen. Für die anderen drei Forschungsansätze werden je drei Studien ins Sample aufgenommen, sodass letztlich neun Publikationen für die Analyse ausgewählt werden.

Das Sampling folgt einem mehrstufigen Stichprobenverfahren, welches sowohl Zufallselemente aufweist, als auch mit bewussten, theoretisch geleiteten, Auswahlentscheidungen vorgeht. Dabei werden die Methoden einer einfachen Zufallsauswahl und die Methoden einer bewussten Auswahl, bzw. Quoten Auswahl nacheinander abgearbeitet.

Auswahlverfahren

Wie oben bereits beschrieben, wurden aus den 49.536 Publikationen der Grundgesamtheit nur diejenigen Publikationen für die Klassifizierung von Zitationsfunktionen mit Erscheinungsjahr 2018 berücksichtigt. Die Einschränkung des Publikationsjahres auf 2018 für die Klassifikation epistemischer Funktionen von Zitierungen (Phase 1) wurde getroffen, um Publikationen mit möglichst vielen Zitierungen, die in den Felddatensatz hinein verweisen zu erfassen. Bei älteren Publikationen steigt die Wahrscheinlichkeit, dass auf Publikationen verwiesen wird, die selbst nicht im Felddatensatz enthalten sind. Da bei der Auswertung die zitationsbasierten thematischen Clusterzuordnungen (aus Leiden7 und Leiden15) berücksichtigt werden sollen, ist dies aber gerade wünschenswert, sodass eine Einschränkung auf das jüngste, vollständig erhobene, Erscheinungsjahr (2018) sinnvoll erschien.

Für das weitere Zusammenstellen des Samples dienten also die 4.396 im Jahr 2018 erschienenen Publikationen als Grundlage. In einem iterativen Verfahren wurde stets eine Publikation zufällig gezogen und auf ihre Eignung für die Aufnahme ins Sample geprüft. Es wurden so lange Ziehungen durchgeführt bis das Sample-Ziel (9 Publikationen, davon jeweils 3 der oben definierten Forschungsansätze: empirisch, Modellierung, anwendungsorientiert) erreicht wurde. Es wurden insgesamt 60 Ziehungen für die Erreichung des Sample-Ziels benötigt. Bei den 60 Ziehungen wurde eine Publikation ein zweites Mal gezogen, sodass es insgesamt 59 verschiedene Publikationen waren, die überprüft wurden.

Die zufällig gezogenen Publikationen wurden unter Berücksichtigung von Titel, Schlagwörtern, Abstract, in unklaren oder schwierigen Fällen auch unter Hinzunahme des Volltextes inhaltlich geprüft. Dazu wurde *erstens* überprüft, ob es sich um eine invasionsbiologische Publikation, eine Publikation mit losem invasionsbiologischen Bezug („invasion-aware“) oder eine Publikation handelt, die keinen invasionsbiologischen Bezug hat. Von den (n=) 59 verschiedenen Publikationen wurden 25 als eindeutig invasionsbiologisch, 25 als invasion-aware und 9 als nicht invasionsbiologisch klassifiziert (siehe 3.1).

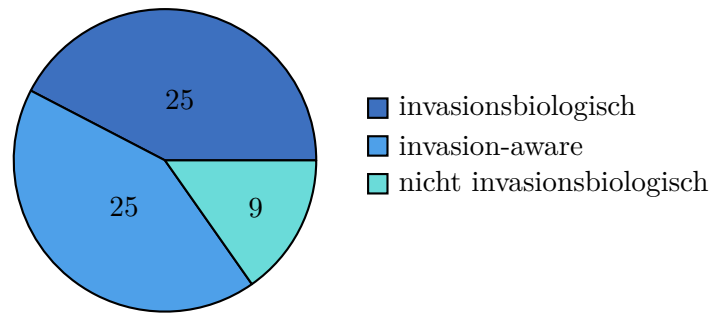


Abbildung 3.1.: Zuordnung der zufällig gezogenen Publikationen hinsichtlich Invasionsbezug

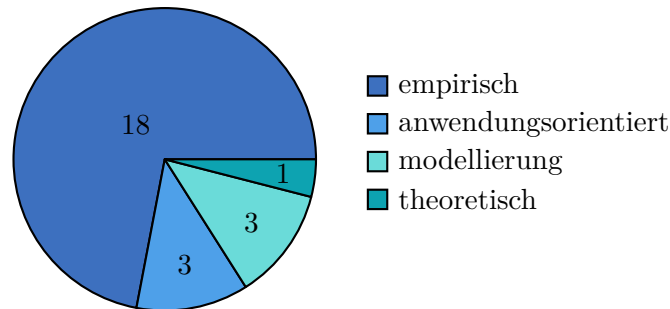


Abbildung 3.2.: Zuordnung der invasionsbiologischen Publikationen hinsichtlich Forschungsansatz

Zweitens wurden diejenigen Publikationen, die eindeutig als invasionsbiologisch einzuschätzen sind ($n=25$) einem der vier Forschungsansätze zugeordnet. Nach inhaltlicher Prüfung der Publikationen konnte bestimmt werden, ob es sich bei dem vorliegenden Beitrag um eine empirische Studie (auf Mikroebene), eine Modellierungsstudie (auf Makroebene), eine anwendungsorientierte oder eine theoretische Arbeit handelt. Von den 25 als eindeutig invasionsbiologisch eingeschätzten Publikationen konnten in erster Instanz 18 dem Typ „empirisch“, vier dem Typ „anwendungsorientiert“ und drei Studien dem Typ „Modellierung“ zugeordnet werden (siehe Abschnitt 3.2). Zunächst wurden keine Publikationen identifiziert, die sich auf rein theoretische Auseinandersetzungen mit invasionsbiologischen Fragestellungen bezogen. Während der Kodierung und damit auch der intensiveren Auseinandersetzung mit den Artikeln ist letztlich ein Artikel aufgefallen, dessen Zuordnung nicht angemessen erschien. Eine Publikationen, die als „anwendungsorientiert“ in das Sample aufgenommen wurde, stellte sich bei genauerer Durchsicht als theoretischer Artikel heraus. Der Artikel wurde deshalb nachträglich aus dem Sample ausgeschlossen und durch eine andere anwendungsorientierte Publikation ersetzt. Letztlich waren von den 25 als eindeutig invasionsbiologisch klassifizierten Publikationen also 18 vom Typ „empirisch“, drei vom Typ „anwendungsorientiert“, drei vom Typ „Modellierung“ und eine vom Typ „theoretisch“.

3.1.2. Sampling Phase 2 – Netzwerkanalyse

Sampling-Ziel

Ziel dieses zweiten Samplings war es, eine Publikation als Seed Dokument für die Netzwerkanalyse auszuwählen. Es soll exploriert werden, welche Implikationen unterschiedliche Zitationsfunktionen auf das bibliometrische Mapping thematischer Strukturen haben.

Für die Auswahl der Publikation für den netzwerkanalytischen Teil spielten verschiedene Überlegungen eine Rolle. Es sollte ein Artikel ausgewählt werden, der recht typisch für das Feld der Invasionsbiologie und im Felddatensatz hinreichend vernetzt ist. Um einen Artikel auszuwählen, dessen eingehende wie auch ausgehende Zitationen möglichst gut im Felddatensatz abgedeckt sind, wurden nur Artikel aus der Mitte des im Felddatensatz abgedeckten Zeitraums für das Sampling berücksichtigt. Es wurden hierfür die Publikationen mit Erscheinungsjahr 2009 ausgewählt. Darüber hinaus sollte es sich um einen „normalen“ Zeitschriftenartikel handeln, nicht etwa um ein Kapitel aus einem Sammelband oder einen Review Artikel. Die ausgewählte Publikation sollte sowohl durchschnittlich zitiert sein, als auch selbst eine durchschnittliche Anzahl an Quellen enthalten. Des Weiteren wurden neben diesen formalen Kriterien auch inhaltliche Aspekte bei der Auswahl berücksichtigt. Es sollte eine Publikation ausgewählt werden, die einen eindeutigen invasionsbiologischen Bezug aufweist und darüber hinaus einen empirischen Ansatz verfolgt. Wie im ersten Sampling-Prozess belegt werden konnte, ist ein überwiegender Teil der klassisch invasionsbiologischen Studien vom Typ „empirisch“, weshalb diese Einschränkung hier sinnvoll erscheint.

Zusammenfassend soll so sichergestellt werden, dass eine für das Feld der Invasionsbiologie typische Publikation für den netzwerkanalytischen Teil ausgewählt wird. Bei dem ausgewählten Paper handelt es sich demnach um ein

- invasionsbiologisches,
- empirisches,
- durchschnittlich zitiertes,
- mit einer durchschnittlichen Anzahl zitierter Referenzen versehenes Paper
- vom Typ „article“ oder „article, proceedings paper“
- aus dem Jahr 2009,
- das im Felddatensatz enthalten ist.

Auswahlverfahren

Die Auswahl des Seed Dokuments für die netzwerkanalytische Auswertung der Implikationen der verschiedenen epistemischen Funktionen von Zitationsbeziehungen folgt einem mehrstufigen Auswahlprozess, der sowohl theoretische Aspekte berücksichtigt als auch Zufallselemente enthält. Für die Auswahl wurden zunächst die 49.536 Publikationen der Grundgesamtheit auf diejenigen beschränkt, die im Jahr 2009 erschienen sind ($n=2.503$). Des Weiteren wurden für das Sampling nur Zeitschriftenartikel und Proceedings Paper berücksichtigt. Nicht für das Sample in Betracht kamen z.B. Review Artikel, da diese üblicherweise besonders oft zitiert sind, oder Kapitel aus Sammelbänden, die üblicherweise besonders selten zitiert sind. Diese Einschränkung reduzierte den Pool der Publikationen aus dem gezogen wurde auf 2.321 Publikationen. Um auch andere Artikel aus dem Sample auszuschließen, die eine besondere Zitationsstruktur aufweisen, wurde für die 2.321 Zeitschriftenartikel und Proceedings Paper die mittleren Zitationsbereiche berechnet. Alle Publikationen mit einer Zitationszahl im Bereich der mittleren 20% (Werte oberhalb des zweiten und unterhalb des dritten Quintils) wurden für die weitere Auswahl berücksichtigt, alle anderen wurden ausgeschlossen. Für die Anzahl der in den Publikationen enthaltenen Referenzen wurde analog verfahren. Konkret bedeutete dies, dass alle Publikationen mit Anzahl eingehender Zitierungen von unter 15 oder über 25 und/oder Anzahl zitierter Referenzen unter 41 und über 51 vom Sample ausgeschlossen wurden. So kamen im Weiteren nur noch 125 Publikationen für die Auswahl in Betracht.

Von diesen 125 Publikationen wurden zufällig Publikationen ausgewählt. Die zufällig ausgewählten Publikationen wurden dann auf ihre inhaltliche Eignung geprüft (analog zur Vorgehensweise im Sampling-

Prozess von Phase 1). Zunächst wurde dabei geprüft, ob es sich bei den jeweils gezogenen Artikeln um klassisch invasionsbiologische Arbeiten handelt. Publikationen, die eher vom Typ „invasion aware“ waren, oder solche bei denen kaum oder gar kein invasionsbiologischer Bezug erkenntlich war, wurden vom Sample ausgeschlossen (siehe 3.1.1). Weil es Ziel des Samplings war, einen empirischen Artikel auszuwählen, wurde geprüft, welchem der vier Typen invasionsbiologischer Publikationen der jeweils vorliegende Artikel zugeordnet werden kann (empirisch, Modellierung, anwendungsorientiert, theoretisch).

Von den 20 zufällig gezogenen Publikationen kamen nach inhaltlicher Prüfung 5 für die Auswahl in Frage. Nach Reihenfolge der Ziehung wurde für diese Publikationen noch geprüft, ob die Abdeckung der zitierten Referenzen und der den Artikel zitierenden Quellen in Web of Science und dem Felddatensatz ausreichend abgedeckt sind. Die erste der fünf empirischen Arbeiten wurde aufgrund dieses Kriterium ausgeschlossen. Von den 50 in der Arbeit zitierten Referenzen waren lediglich 9 in Web of Science enthalten. Die zweite der fünf Publikationen wies eine gute Abdeckung der zitierten und zitierenden Quellen (35 von 42 zitierten Quellen und 25 von 25 zitierenden Quellen) in Web of Science auf, von denen wiederum auch eine angemessene Anzahl im Felddatensatz enthalten ist (12 von 35 in Web of Science enthaltenen zitierten Quellen und 22 von 25 zitierenden Quellen). Die Publikation entsprach somit allen definierten Kriterien und wurde für das Sample ausgewählt.

3.2. Klassifikation

Unabhängig davon ob nun von Denkstilen, anerkannten Problemlösungen im Rahmen normalwissenschaftlicher Forschung, Wissenskulturen oder epistemischen Regimen die Rede ist, ist über die in Kapitel 2 in aller Kürze skizzierten Ansätze hinweg die Erkenntnis festzuhalten, dass wissenschaftliche Forschung sich innerhalb wie auch immer bezeichneter Teil-Gemeinschaften organisiert und dass innerhalb dieser Teil-Gemeinschaften je spezifische, die Forschungsarbeit strukturierende, Überzeugungen und Praktiken geteilt werden. Diese Überzeugungen und Praktiken hinterlassen neben flüchtigen auch manifeste Spuren, die sich einer rekonstruktiven Analyse unterziehen lassen. Die Rekonstruktion von Sinnstrukturen in manifesten Spuren menschlichen Handelns, oder im vorliegenden Fall, wissenschaftlichen Forschungshandelns, ist Ziel der vorliegenden Arbeit. Zitationen in wissenschaftlichen Publikationen sind solche manifesten Spuren wissenschaftlichen Forschungshandelns und öffnen sich als solche der Rekonstruktion von in wissenschaftlichen Teil-Gemeinschaften geteilten, die Forschungsarbeit strukturierenden Überzeugungen und Praktiken.

Konzeptionelle Überlegungen zum Kodieren

Der Analyse von Zitationsbeziehungen gehen einige konzeptionelle Überlegungen voraus. Wissenschaftliche Publikationen beinhalten in der Regel vielfältige Wissensbeiträge (sog. „knowledge claims“). In einer die Publikation zitierenden Arbeit kann daher Bezug auf ganz verschiedene epistemische Funktionen genommen werden. Aus dieser Überlegung ergibt sich, dass Zitationsbeziehungen zwischen zwei Publikationen, also dem zitierten und dem zitierenden Dokument, nicht immer die selbe epistemische Funktion ausdrücken. Die Einbettung des Zitats in der zitierenden Publikation wird dadurch besonders relevant. Methodisch ergibt sich daraus für die Kodierung epistemischer Funktionen von Zitationsbeziehungen die Notwendigkeit der Berücksichtigung sowohl des Inhalts der zitierten Publikation im Sinne einer Inhaltsanalyse als auch die Berücksichtigung des Kontextes des Zitats in der zitierenden Publikation im Sinne einer Kontextanalyse. Für die Bestimmung der Zitationsbeziehungen wurden daher sowohl Kontextinformationen in den zitierenden Publikationen als auch die Inhalte der zitierten

Publikationen miteinbezogen. Die Metadaten zu dem zitierten Text wie Titel, Schlagwörter, Abstract und dergleichen, sowie Platzierung des Zitates in dem zitierenden Text haben die Interpretation unterstützt.

Aus diesen Überlegungen ergibt sich auch, dass jede zitierte Quelle an der Stelle ihrer Zitation beurteilt werden muss. Einer Referenz im Literaturverzeichnis einer Publikation entsprechen unter Umständen mehrere Zitate im Text. Jeder dieser Zitatstellen kann (und muss m.E.) ein epistemischer Bezug zugeordnet werden. Dieses Vorgehen ermöglicht die Berücksichtigung der Tatsache, dass ein und dieselbe Publikation an verschiedenen Stellen für unterschiedliche epistemische Funktionen zitiert wird. Zugleich ist es möglich, dass Zitationen mehr als eine funktionelle Beziehung gleichzeitig verkörpern, weshalb prinzipiell die Möglichkeit der doppelten Kodierung epistemischer Funktionen zugelassen wurde.

Entwicklung des Kategoriensystems

Zitationsbeziehungen können ganz unterschiedliche thematische Beziehungen zwischen zitierender Publikation und zitierter Quelle ausdrücken. So können Zitate z.B. auf die Nachnutzung einer bestimmten Methode, spezifische Eigenschaften des Untersuchungsgegenstandes, den Forschungsstand zum bearbeiteten Forschungsproblem, verschiedene Anwendungsszenarien oder grundlegende theoretische Überlegungen verweisen. Der Forschungsstand zur Analyse von Zitationsfunktionen (vgl. Abschnitt 2.3) unterstreicht die Vielfalt der möglichen Beziehungen. Für die Analyse wurde ein Kategoriensystem entwickelt, das einige Elemente vorheriger Studien aufgreift. Daneben gingen praktische Überlegungen in die Entwicklung des Kategoriensystems mit ein.

Es sollte ein Kategoriensystem entwickelt werden, welches der Verschiedenheit epistemischer Beziehungen hinreichend gerecht wird, gleichzeitig aber nicht übermäßig differenziert ist. Ein sehr umfangreiches Kategoriensystem erschien hinsichtlich der weiteren Verwendung der Kategorien als Grundlage für die Erstellung multiplexer Netzwerke und hinsichtlich der deskriptiven Analyse der neun ausgewählten Publikationen nicht praktikabel⁵. Für die vergleichende Analyse wissenschaftlicher Felder als epistemische Regime schlagen Gläser u. a. (2018) einige relevante Aspekte vor. Wissenschaftliche Felder lassen sich nach den Autoren u.a. hinsichtlich der verwendeten Methoden, der untersuchten Gegenstände, des bearbeiteten Forschungsproblems, der Bezugnahme auf praktische Anwendungskontexte und der theoretischen Ausrichtung unterscheiden und vergleichen. Aus diesen konkreten Überlegungen wurden die fünf Kategorien für das Kategoriensystem der vorliegenden Arbeit abgeleitet: research problem, empirical object, method, theoretical, application.

Zitate, die einen Bezug auf das Forschungsproblem herstellen, werden mit dem Kode „research problem“ versehen. Diese Bezüge betreffen die grundlegende Fragestellung der Arbeit und sind daher sehr individuell für jede Publikation zu bestimmen. Innerhalb der Invasionsbiologie betrifft das Forschungsproblem häufig das Auftreten und die Verbreitung invasiver Arten, die Mechanismen von Invasionsprozessen inklusive deren Faktoren für Erfolg und Misserfolg, aber auch die Möglichkeiten zur Bekämpfung bestimmter schädlicher invasiver Arten. Der Kode „empirical object“ drückt Bezüge aus, die den empirischen Gegenstand behandeln. Die empirischen Objekte invasionsbiologischer Studien sind zumeist bestimmte, botanische oder zoologische, invasive Arten. In biologischen Studien ist es üblich bei Namensnennung einer bestimmten, bereits klassifizierten Art den Erst-Entdecker der Art zu

⁵Da sich die Codes als lediglich nominal skalierte Variablen interpretieren lassen, lag eine Analyse der Daten in Form von Kreuztabellen und entsprechenden Chi-Quadrat-Signifikanztests auf der Hand. Für diese Art der Auswertung ist die ausreichende Besetzung aller Zellen Bedingung und lässt sich durch eine geringere Anzahl an Kategorien eher verwirklichen.

zitieren. Zitierungen dieser Art betreffen direkt den Untersuchungsgegenstand. Darüber hinaus sind auch Beschreibungen der Eigenschaften bestimmter Arten sowie Beschreibungen deren Lebensraums mit dem Kode „empirical object“ zu verstehen. Der Kode „method“ drückt Bezüge aus, bei denen es um Verweise auf die Verwendung bestimmter Methoden, Modelle oder Formeln geht. Üblicherweise wird in Zeitschriftenartikeln das methodische Vorgehen zur intersubjektiven Nachvollziehbarkeit eingehend beschrieben. Dabei wird häufig Bezug auf Methoden genommen, die sich bereits im Feld etabliert haben, die kritisiert werden oder die sich noch als valide erweisen müssen. In der Invasionsbiologie gibt es eine weitreichende Methodenpluralität, die von Feldbeobachtungen und bildgebenden Verfahren über Experimente und Gensequenzanalysen zu statistischen Modellen reicht. Oft wird bei der Beschreibung des eigenen methodischen Vorgehens auf in anderen Studien verwendete oder entwickelte Verfahren, Methoden, Modelle oder Formeln in Form von Zitaten verwiesen. Auch Bezüge zu Samplingstrategien und die Beschreibung von Schwachstellen betreffen in diesem Sinne die Methode. Der Kode „theoretical“ drückt Bezüge aus, die auf Theorien verweisen. Theoretische Referenzen verweisen in invasionsbiologischen Studien häufig auf generelle Eigenschaften, Abläufe und Zusammenhänge von Invasionsprozessen. Auch einige konkrete Hypothesen lassen sich aus diesen allgemeinen Theorien ableiten und können daher ebenfalls als theoretische Bezüge verstanden werden. Der Kode „application“ dient der Klassifizierung von Referenzen, die auf konkrete Anwendungskontexte verweisen. Diese Anwendungskontexte sind für Governance-Prozesse besonders relevant. Es geht dabei häufig um negative Auswirkungen bestimmter invasiver Arten oder um konkrete Management-Maßnahmen, die der weiteren Verbreitung der Art Einhalt gebieten soll. Häufig werden auch politische Programme erwähnt, die konkrete Richtlinien, Schwellenwerte oder Zielvorgaben definieren.

Die fünf Subkodes wurden deduktiv gebildet. Es wurde dabei offen gehalten während des Kodierungsprozesses dem Kategoriensystem bei Bedarf weitere Kategorien (induktiv) hinzuzufügen. Bei der Arbeit in den Publikationen stellte sich das Kategoriensystem aber als sehr praktikabel und ausreichend differenziert heraus, sodass keine weiteren Kategorien hinzugefügt wurden.

Neben der epistemischen Funktion wurden noch zwei weitere Dimensionen kodiert: der Ort des kodierten Segments innerhalb der Publikation nach Kapiteln und ob es sich um eine Selbstzitation handelt oder nicht. Beide Dimensionen wurden auch schon in anderen Studien untersucht (vgl. für Selbstzitationen etwa Garfield, 1978, Garfield und Welljams-Dorof, 1990, Kochan und Budd, 1992 und für die Kodierung des Zitationsortes etwa Maričić u. a., 1998). Für die deskriptive Analyse der Zitationsbeziehungen und deren Verteilung wurden die Dimensionen als besonders relevant eingeschätzt. Für die Kodierung des Ortes der Zitation wurden zwei Kodes mit jeweils mehreren Subkodes gebildet. Der Kode „location_orig“ enthält als Subkodes die induktiv erhobenen Ortsangaben, die sich aus den Kapitelüberschriften im Text ergeben. Der Kode „location_standardized“ enthält als Subkodes vier deduktiv gebildete Kategorien, die die typische Gliederung von Zeitschriftenartikeln widerspiegeln: „introduction“, „methods (incl. materials)“, „results“, „discussion (incl. conclusion)“. Für die Vergabe des Kodes „self-citation / no self-citation“ wurden die Autorenlisten der zitierenden und der zitierten Publikation durchgesehen und auf Übereinstimmung geprüft. Die Segmente wurden mit „self-citation“ kodiert, sobald auch nur ein Autor oder eine Autorin an beiden Publikationen beteiligt war. Außerdem wurde für jede zitierte Referenz die WoS-ID (Accession Number, UT) als Kommentar dem kodierten Segment beigelegt. Dadurch wurde eine spätere Zusammenführung mit Metadaten aus dem Web of Science und aus dem Forschungsprojekt von Frau Velden ermöglicht.

Vorgehensweise und Herausforderungen beim Kodieren

Für die Kodierung wurde mit dem qualitativen Datenanalyseprogramm MaxQDA 2020 (*MAXQDA 2020: Software für qualitative Datenanalyse*) gearbeitet. Für jeden der neun ausgewählten Artikel wurden die Volltexte im Portable Document Format recherchiert und in MaxQDA hinterlegt. Dadurch wurde die Kodierung der zitierten Segmente direkt im Volltext ermöglicht (siehe Abbildung 3.3).

Für eine erste Einschätzung der inhaltlichen Ausrichtung der Publikationen und die Festsetzung konkreter Anhaltspunkte für die Kodierung der epistemischen Funktionen wurden alle Artikel zunächst durchgelesen. Darauf aufbauend wurden dann zusammenfassende Synopsen für jedes Dokument erstellt (siehe A.2 Übersicht Sample Invasionsbiologie, S. 71). Diese Synopsen wurden in Form von Dokumentenmemos den Volltexten angeheftet und enthielten jeweils eine kurze Inhaltsbeschreibung und die Einschätzung möglicher Inhalte für die Kodierung der epistemischen Funktionen. Die Synopsen wurden im Verlauf des iterativen Kodierungsprozesses mehrfach überarbeitet, korrigiert und präzisiert. Sie dienten als erste Orientierung bei der Kodierung und nahmen im Verlauf des Kodierungsprozesses die Rolle eines groben Kodierleitfadens ein. Es wurde jeweils definiert wie das Forschungsproblem der je vorliegenden Publikation definiert wurde, welche Theorien, Methoden, empirischen Objekte und Anwendungszusammenhänge in den Artikeln besprochen wurden und wie diese jeweils aufgefasst und voneinander abgegrenzt wurden.

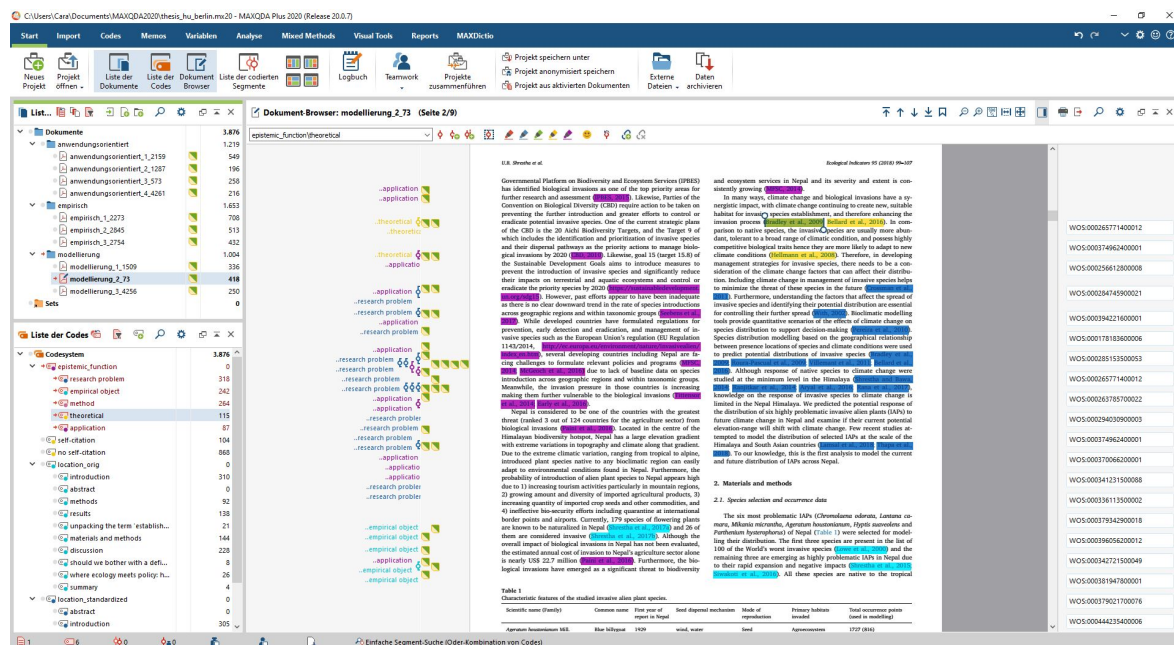


Abbildung 3.3.: Screenshot Kodierung mit dem Datenanalyseprogramm MaxQDA

Nach dem ersten Durchlesen und dem Anfertigen der Synopsen wurden alle zu kodierenden Segmente markiert und der Reihe nach kodiert. Dazu wurde sowohl der Kontext in dem die Segmente eingebettet sind berücksichtigt, als auch der Inhalt der zitierten Artikel. Jede Quelle wurde dazu in Web of Science recherchiert, wo zumeist Informationen wie Titel, Schlagwörter und Abstract die Interpretation der epistemischen Funktion erleichterten. Quellen, die nicht in WoS verzeichnet waren, wurden über andere Literatursuchmaschinen (vornehmlich Google Scholar) recherchiert. Alle kodierten Segmente wurden darüber hinaus mit Memos versehen. In diesen Memos wurden z.B. Anmerkungen zur Kodierentscheidung, Inhalte der zitierten Studie und die Schwierigkeit der Einschätzung dokumentiert. Die Dokumentation der Schwierigkeit der Einschätzung diente sowohl der Reflektion des

Kodierungsprozesses als auch als Anhaltspunkt für die erneute Überarbeitung und Überprüfung der Interpretation.

Alle Dokumente wurden in einem iterativen Verfahren mehrfach durchgesehen, alle kodierten Segmente mehrfach geprüft und deren Richtigkeit und Konsistenz beurteilt. An einigen Stellen wurde bei erneuter Durchsicht die Kodierung geändert oder eine Doppelkodierung von epistemischer Funktion hinzugefügt. Die intensive Auseinandersetzung mit dem Material und seinen Inhalten führte zu Lerneffekten und Routinisierung der Kodiertätigkeit. Auch wenn keine systematische Erhebung der Intrakoder-Reliabilität erfolgte, wurde durch das iterative Verfahren, die stetige kritische Auseinandersetzung mit den eigenen Interpretationen und die andauernde Überarbeitung der Synopsen, die als Kodierleitfaden dienten, eine konsistente Kodierung gewährleistet.

Nach der Erstellung der zusammenfassenden Synopsen und der Kodierung des ersten Dokuments wurde eine Interpretationssitzung zur Zuordnung der Subkodes für epistemischen Funktionen mit Frau Velden durchgeführt, die durch ihre Forschung u.a. über das Feld der Invasionsbiologie tiefergehende Kenntnisse und Erfahrungen mit invasionsbiologischen Publikationen mitbringt. In dieser Sitzung wurde über die zusammenfassenden Synopsen und die darin jeweils definierten Inhalte der Subkodes diskutiert und es wurde die Zuordnung einzelner Segmente ausschnittsweise abgeglichen. Die Diskussion ergab, dass die Inhalte der Subkodes nahezu deckungsgleich definiert und interpretiert wurden und auch der Abgleich einzelner (je unabhängig voneinander kodierten) Segmente ergab eine hohe Übereinstimmung. Zwar kann dies kein Ersatz für eine systematische Testung von Interkoder-Reliabilität sein, dennoch zeigt die hohe Übereinstimmung die Konsistenz und Nachvollziehbarkeit des Codesystems. Die Kodierung der anderen Subkodes (`self-citation` `no self-citation` und `location_orig` und `location_standardized`) und die Kommentierung mit den (wenn vorhanden) entsprechenden WoS-IDs erfolgte parallel.

Erstellung des Datensatzes

Nach der Fertigstellung der Kodierung wurden die Informationen aus MaxQDA exportiert, um sie für die Zusammenführung mit den Informationen aus dem Felddatensatz vorzubereiten. Außerdem wurden die Informationen noch um einige Metadaten aus dem Web of Science ergänzt. Da die aus MaxQDA exportierten Daten nicht in einer Form vorlagen, die für die weitere Verarbeitung praktikabel war, wurden die Daten zunächst in Microsoft Excel und weiter in Open Refine aufbereitet⁶. Nach der Aufbereitung wurde die in Tabellen vorliegenden Daten mit einigen Metadaten aus dem Web of Science und den Informationen aus dem Felddatensatz zusammengeführt. Die Informationen aus dem Felddatensatz, bzw. zu den Zuordnungen zu den Clusterlösungen stellte Frau Velden zur Verfügung. Über die WoS-IDs (Accession Number, UTs) konnten die Informationen in Microsoft Excel unter Nutzung der Funktion SVERWEIS gematcht werden. Analog wurde für die Anreicherung des Datensatzes mit Metadaten aus dem Web of Science verfahren, die zunächst über die Export Funktion der Weboberfläche exportiert wurden.

Insgesamt ergab sich ein Datensatz mit 24 Variablen. Die Hälfte davon stammt aus der Kodierung in MaxQDA, weitere sechs aus Web of Science und die verbleibenden sechs aus dem Forschungsprojekt von Frau Velden. Die Daten wurden anschließend für Auswertung in R CRAN (R Core Team, 2017) eingelesen.

⁶Insbesondere die doppelt kodierten Stellen erschwerten die Datenaufbereitung. Sie mussten zunächst aus dem Datensatz herausgelöst, getrennt aufbereitet und später wieder hinzugefügt werden. Die Trennung der doppelt und der einfach kodierten Segmente wurde in Microsoft Excel umgesetzt. Die weitere Aufbereitung wurde z.T. auch in Open Refine durchgeführt.

3.3. Netzwerkmodellierung

Für den netzwerkanalytischen Teil der vorliegenden Arbeit wurde ein Zeitschriftenartikel ausgewählt, der als „Seed Dokument“ den Kern der Netzwerke bildet (zur Auswahl des Seed Dokuments siehe Kapitel 3.1.2, S. 25). Ausgehend von diesem Seed Dokument (SD) wurden in Web of Science alle den Artikel zitierenden und alle in dem Artikel zitierten Quellen ermittelt. Im Seed Dokument werden 12 Artikel zitiert, die in WoS verzeichnet sind. Das Seed Dokument wird von 22 Artikeln zitiert, die in WoS verzeichnet sind. Die Metadaten der insgesamt 35 Artikel (SD + 12 zitierte + 22 zitierende Artikel) wurden aus der Datenbank exportiert und stellen die Grundlage für die Erstellung des Primärnetzwerks dar. Für die Erstellung des Netzwerks wurden zusätzlich die in den 12 zitierten Artikel und 22 zitierenden Artikeln enthaltenen Referenzen ermittelt und exportiert. Insgesamt konnten so 400 in Web of Science verzeichnete Quellen ermittelt werden, die entweder direkt oder über einen Verbindungsknoten mit dem Seed Dokument verbunden sind.

Für die Klassifikation der Zitationsfunktionen wurden das Seed Dokument, die 12 darin zitierten Artikel (r1–r12) und die 22 das Seed Dokument zitierenden Artikel (z1–z22) als Volltexte in MaxQDA eingepflegt. Zunächst wurden über einen Abgleich mit den exportierten bibliographischen Daten alle zu kodierenden Stellen ermittelt. Es wurden nur diejenigen Zitate ausgewählt, die auch in Web of Science verzeichnet sind. Um später eine eindeutige Zuordnung zu ermöglichen, wurden die zu kodierenden Textstellen (Zitationen) mit ihrer WoS-ID versehen. Für eine erste Einschätzung der inhaltlichen Ausrichtung der Publikationen und die Festsetzung konkreter Anhaltspunkte für die Kodierung der epistemischen Funktionen wurden alle Artikel zunächst durchgelesen (analog zu Kapitel 3.2, S. 30). Für die Kodierung hilfreiche Anhaltspunkte wurden in Dokumentmemos festgehalten (siehe Abbildung 3.4).⁷ Nach dem Kennzeichnen der zu kodierenden Stellen, der Vergabe der WoS-IDs und einer ersten inhaltlichen Durchsicht der Artikel, inklusive der Anfertigung der Dokumentmemos zur Orientierung bei der Kodierung wurden dann schließlich die ausgewählten Zitierungen kodiert. Für die Kodierung wurde analog zum ersten Teil zwischen research problem, epistemic function, method, theoretical und application unterschieden. Auch hier wurde ein iteratives Verfahren angewendet, alle Kodierungen wurden mehrfach durchgesehen bis eine konsistente und vollständige Kodierung aller Stellen gewährleistet war.

Im Anschluss wurden die generierten Daten aus MaxQDA exportiert und für die weitere Bearbeitung, bzw. dem Erstellen von Netzwerkkarten aufbereitet. Die MaxQDA Exporte enthalten Informationen zur Dokumentgruppe (citing, cited, seed document), zum Dokumentnamen, zum Ort der Zitation (Anfang, Ende), zur kodierten epistemischen Funktion (research problem, empirical object, method, theoretical, application), zum kodierten Segment und die WoS-ID der zitierten Referenz. Dokumentname und WoS-ID sind als Knoten zu interpretieren, die epistemische Funktion sagt etwas über die inhaltliche Art der Verbindung zwischen den Knoten aus. Die WoS-IDs wurden zur einfacheren Interpretation und übersichtlicheren Beschriftung in den Grafiken in selbst gewählte IDs überführt (r1–r12 für die zitierten Referenzen und z1–z22 für die das Seed Dokument zitierenden Artikel; analog zur Bezeichnung der Dokumente). Aus den so aufbereiteten Informationen wurden Tabellen für Knoten und Kanten für den Import in die Netzwerkanalysesoftware Gephi (Bastian, Heymann und Jacomy, 2009) erstellt.

Für die inhaltliche Analyse sollten die in den Publikationen besprochenen Themen (Topics) herausgearbeitet werden. Aufgrund der Größe und zu erwartender Unübersichtlichkeit des Gesamtnetzwerkes

⁷Die Dokumentmemos strukturieren sich in gleicher Form wie die zusammenfassenden Synopsen, die in Phase eins für die Analyse der Zitationsfunktionen angefertigt wurden (siehe dazu auch A.2 Übersicht Sample Invasionsbiologie, S. 71). An dieser Stelle wurde darauf verzichtet alle 35 Dokumentmemos in den Anhang mit aufzunehmen. Die Informationen finden sich allerdings in den MaxQDA Dateien, die auf dem der Arbeit beigelegten elektronischen Datenträger gespeichert sind.

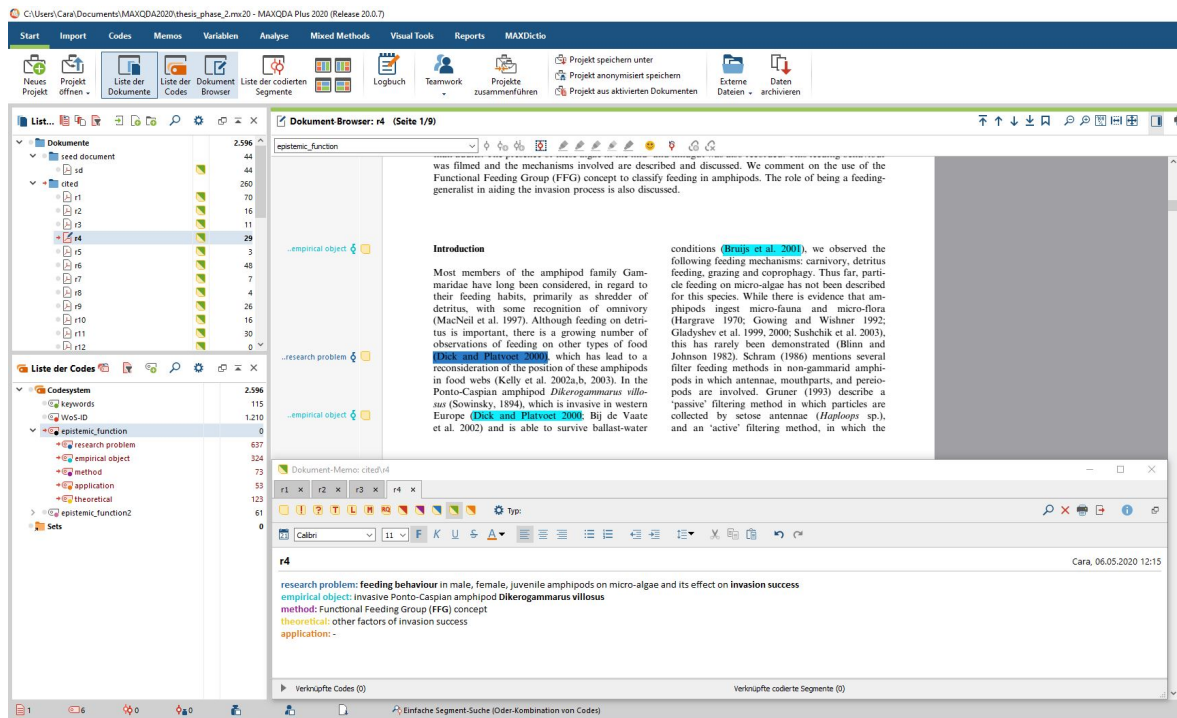


Abbildung 3.4.: Screenshot Kodierung und Dokumentmemos mit dem Datenanalyseprogramm MaxQDA

wurde für die Analyse nur ein Teil des Netzwerkes, das Primärnetzwerk, ausgewählt. Das Primärnetzwerk enthält alle Publikationen, die direkt mit dem Seed Dokument verbunden sind, sowie ihre Verbindungen untereinander. Das Primärnetzwerk enthält damit 35 Knoten und 282 Verbindungen. Sie bilden die Grundlage für die Identifizierung von Topics. Die Topics wurden aus dem jeweiligen Kontext der Zitationen ermittelt. Dazu wurden die entsprechenden Textstellen innerhalb der Volltexte in MaxQDA erneut betrachtet und aus dem Kontext die Inhalte und Themen auf die sich die Zitate beziehen herausgearbeitet. Dabei wurden übergreifende Termini entwickelt, die Gleiches zusammenführen und Verschiedenes trennen. Die ermittelten Topics wurden als Kantenbeschriftungen den Netzwerkgrafiken beigelegt um eine visuelle Analyse zu ermöglichen.

4. Ergebnisse

4.1. Analyse der Klassifizierung

4.1.1. Überblick Datensatz und kodierte Segmente

Für die Analyse wurden insgesamt neun Artikel vollständig kodiert (siehe Tabelle 4.1). Davon sind je drei vom Typ „anwendungsorientiert“, drei vom Typ „empirisch“ und drei vom Typ „Modellierung“ in die Analyse eingegangen. Die Artikel enthalten insgesamt 568 Referenzen, wovon 238 (rund 42%) auf die empirischen Artikel fallen. In den anwendungsorientierten Publikationen werden 170 Referenzen zitiert (30%) und in den Publikationen des Typs „Modellierung“ finden sich die restlichen 160 Zitationen (28%). Da Referenzen häufig mehrfach, an verschiedenen Stellen einer Publikation zitiert werden, ist die Anzahl der kodierten Segmente mit 907 deutlich höher als die Anzahl der Referenzen. Aus dem Verhältnis lässt sich ableiten, dass jede in den Literaturangaben verzeichnete Quelle durchschnittlich rund 1,6 Mal im Text Erwähnung findet. Das Verhältnis von kodierten Segmenten zu im Text enthaltenen, eindeutigen Referenzen unterscheidet sich nach Dokumentengruppe kaum (anwendungsorientiert = 1,5, modellierung = 1,5, empirisch = 1,7). Bei der Kodierung der epistemischen Funktionen von Zitationsbeziehungen wurde eine Doppelkodierung zugelassen, weshalb die Zahl der vergebenen Codes die Anzahl kodierter Segmente übersteigt. Insgesamt wurden 959 Subkodes zur Beschreibung der Zitationsbeziehungen vergeben.

Dokumentname, bzw. Dokumentgruppe	Artikel	Referenzen	kodierte Segmente (exkl. Doppelkodierung)	kodierte Segmente (inkl. Doppelkodierungen)
anwendungsorientiert_1_2159	1	93	135	144
anwendungsorientiert_3_573	1	42	64	66
anwendungsorientiert_4_4261	1	35	24	54
anwendungsorientiert	3	170	253	264
empirisch_1_2273	1	78	177	177
empirisch_2_2845	1	75	127	132
empirisch_3_2754	1	85	108	108
empirisch	3	238	412	417
modellierung_1_1509	1	50	76	108
modellierung_2_73	1	75	104	106
modellierung_3_4256	1	35	62	64
modellierung	3	160	242	278
gesamt	9	568	907	959

Tabelle 4.1.: Übersicht Auswertung Fallzahlen

Von den 907 kodierten Segmenten sind rund 63% in Web of Science verzeichnet. Zu diesen 575 Zitationen konnten zusätzlich zu den in MaxQDA erstellten Variablen weitere Informationen hinzugefügt werden. Dabei unterscheidet sich die Abdeckung der Referenzen in Web of Science zwischen den

verschiedenen Dokumentgruppen in geringem Maße. Die Referenzen aus den Artikeln vom Typ „Modellierung“ sind hierbei mit 67% am besten abgedeckt, Referenzen aus den anwendungsorientierten Artikeln sind zu 64% indexiert und Referenzen aus den empirischen Publikationen sind zu 61% in Web of Science verzeichnet. Aufgrund der kleinen Fallzahl, sollten diese Unterschiede aber nicht überbewertet werden.

4.1.2. Epistemische Funktion

Insgesamt wurden bei der Kodierung der epistemischen Funktionen 959 Subkodes zur Beschreibung der Zitationsbeziehungen vergeben. Mit 264 und 265 kodierten Segmenten treten die Codes „research problem“ und „method“ am Häufigsten auf. Auf die beiden Kategorien fallen jeweils rund 28% der analysierten Zitate. Dritthäufigste Funktion bilden mit rund 25% Zitate, die sich auf empirische Objekte beziehen. Unter der bei statistischer Unabhängigkeit zu erwartenden absoluten Häufigkeit¹ liegen Zitate mit theoretischem Bezug (rund 11%) und Zitate mit Anwendungsbezug (rund 9%). Betrachtet man die Verteilung der Zitationsfunktionen nach Dokumentgruppe getrennt, fallen weitere Unterschiede auf. So scheinen etwa anwendungsbezogene Zitate in anwendungsorientierten Publikationen deutlich überrepräsentiert (60%) und in empirischen Arbeiten deutlich unterrepräsentiert (6%)². Zitate, die sich auf empirische Objekte beziehen, scheinen dahingegen überproportional häufig in empirischen Publikationen vorzukommen (56%), weniger in Publikationen vom Typ „Modellierung“ (15%). Dafür sind methodische Zitationen häufiger als in allen anderen Publikationstypen in Modellierungsartikeln zu finden (43%). In anwendungsorientierten Publikationen sind verglichen damit deutlich weniger Zitationen mit methodischem Bezug zu finden (16%).

Dokumentname, bzw. Dokumentgruppe	application	empirical object	method	research problem	theoretical	summe
anwendungsorientiert_1_2159	16	60	12	31	25	
anwendungsorientiert_3_573	16	2	15	20	13	
anwendungsorientiert_4_4261	17	6	15	14	2	
anwendungsorientiert	49	68	42	65	40	264
Zeilen-%	19%	26%	16%	25%	15%	101%
Spalten-%	60%	28%	16%	25%	37%	
empirisch_1_2273	2	93	0	75	7	
empirisch_2_2845	2	15	73	22	20	
empirisch_3_2754	1	26	36	24	21	
empirisch	5	134	109	121	48	417
Zeilen-%	1%	32%	26%	29%	12%	100%
Spalten-%	6%	56%	41%	46%	44%	
modellierung_1_1509	4	68	68	32	2	
modellierung_2_73	24	32	32	19	15	
modellierung_3_4256	10	13	12	28	4	
modellierung	28	37	113	79	21	278
Zeilen-%	10%	13%	41%	28%	8%	100%
Spalten-%	34%	15%	43%	30%	19%	
gesamt	82	239	264	265	109	959
Zeilen-%	9%	25%	28%	28%	11%	101%
Spalten-%	100%	99%	100%	101%	100%	

Tabelle 4.2.: Übersicht absolute Häufigkeiten: epistemic_function

¹Absolute erwartete Häufigkeiten lassen sich in Kreuztabellen über das Verhältnis des Produkts aus Spalten- und Zeilensummen zur Gesamtanzahl der Beobachtungen (N) berechnen (Spaltensumme mal Zeilensumme geteilt durch N). Siehe Anhang B.1, S. 81).

²Siehe hierzu die Spaltenprozentage für Zitate mit der Funktion „application“ in Tabelle 4.2.

Insgesamt legen diese Beobachtungen nahe, dass die Verteilung von Zitationsfunktionen nach Dokumentgruppe variiert. Vergleicht man die Spaltenprozentage mit der bei statistischer Unabhängigkeit zu erwartenden Verteilung dieser Häufigkeiten, wird der Zusammenhang besonders deutlich. Ein Chi²-Unabhängigkeitstest über alle epistemischen Funktionen und alle drei Dokumentgruppen bestätigt die Vermutung: Es zeigt sich ein höchst signifikanter Zusammenhang zwischen den beiden Eigenschaften³. Mit Cramer's V = 0,25075 lässt sich der Zusammenhang als mittelstark interpretieren. Betrachtet man den Zusammenhang je einzeln für die fünf Kategorien der Variable „epistemic_function“, zeigen sich schwache bis mittlere, jedoch höchst signifikante ($p < 0,0001$) Zusammenhänge für Zitate vom Typ „application“ (Cramer's V = 0,26), „empirical object“ (Cramer's V = 0,18) und „method“ (Cramer's V = 0,21). Für Zitate mit Bezug auf das Forschungsproblem und Zitate, die auf einen theoretischen Kontext verweisen, konnte kein Zusammenhang nachgewiesen werden. Theoretische und das Forschungsproblem betreffende Zitationen sind demnach in allen drei Dokumentgruppen gleichermaßen vertreten. Dahingegen unterschieden sich Typen invasionsbiologischer Publikationen hinsichtlich der Verwendung anwendungsbezogener, methodischer und auf das empirische Objekt bezogener Quellen.

4.1.3. Ort der Zitierung

Um einen Überblick darüber zu bekommen wie sich die Zitationen auf die verschiedenen Abschnitte eines Aufsatzes verteilen und ob sich diese nach epistemischer Funktion unterscheiden, wurde der Ort der Zitation ebenfalls kodiert (siehe Tabelle 4.3). Von den 907 kodierten Segmenten befinden sich rund ein Drittel (34%) in den jeweiligen Einleitungskapiteln. Ein Großteil der restlichen Quellen fällt zu ungefähr gleichen Teilen auf die Methodenkapitel (26%) und Diskussionsteile (inklusive Fazit/Konklusion, 25%). In Kapiteln, in denen die Ergebnisse präsentiert und erklärt werden, wird kaum zitiert.

Dokumentname, bzw. Dokumentgruppe	introduction	methods (incl. materials)	results	discussion (incl. conclusion)	summe
anwendungsorientiert_1_2159	66	31	0	38	
anwendungsorientiert_3_573	27	14	0	23	
anwendungsorientiert_4_4261	25	12	3	14	
anwendungsorientiert	118	57	3	75	253
Zeilen-%	47%	23%	1%	30%	
empirisch_1_2273	30	0	127	20	
empirisch_2_2845	26	48	8	45	
empirisch_3_2754	31	42	0	35	
empirisch	87	90	135	100	412
Zeilen-%	21%	23%	33%	24%	
modellierung_1_1509	24	40	0	12	
modellierung_2_73	41	33	0	30	
modellierung_3_4256	35	16	0	11	
modellierung	100	89	0	53	242
Zeilen-%	41%	37%	0%	22%	
gesamt	305	236	138	228	907
Zeilen-%	34%	26%	15%	25%	100%

Tabelle 4.3.: Übersicht absolute Häufigkeiten: location_stand

Eine einzige Publikation im Sample (empirisch_1_2273) verzerrt diese generelle Tendenz um ein beachtliches Maß. In der Studie werden 127 der insgesamt 138 auf Ergebniskapitel entfallenden Zitationen getätigt. Das entspricht 92% der in Ergebniskapiteln zitierten Quellen und 72% der in empirisch_1_2273 zitierten Quellen. Schließt man diese Publikation von der Berechnung der relativen Anteile aus, entfallen nur 1,4% der insgesamt zitierten Quellen auf Ergebniskapitel. Auffällig ist für

³mit $\text{Chi}^2 = 120,6$, $\text{df} = 8$, $p < 2,2e^{-16}$.

den Artikel darüber hinaus, dass keine einzige Quelle im Methodenabschnitt zitiert wird, was ebenfalls deutlich von der generellen Tendenz abweicht. Aufgrund der geringen Größe des Samples ist nicht abschließend zu beurteilen, ob die Publikation „empirisch_1_2273“ einen echten Ausreißer darstellt oder ob sie etwa einen eigenen Typus hinsichtlich ihrer Zitationsstruktur bildet. Abgesehen von empirisch_1_2273 zeigt die Verteilung der Zitate auf die verschiedenen Kapitel keine nennenswerten Unterschiede zwischen Publikationen und Publikationsgruppen.⁴

Neben der generellen Beschreibung der örtlichen Zitationsstruktur (wo wird zitiert?) ist es interessant diese hinsichtlich der zitierten Inhalte näher zu betrachten (was wird wo zitiert?). Dazu wurden für alle vergebenen Codes zur epistemischen Funktion der Zitate die jeweiligen Kapitel ermittelt, in denen die Quellen zitiert wurden. Zitate, die zwei verschiedene epistemischen Funktionen einnehmen, sind demnach in der Analyse doppelt berücksichtigt (n=959). Zur Visualisierung dieses Zusammenhangs dienen Abbildung 4.1, die die prozentuale Verteilung der in den jeweiligen Kapiteln enthaltenen Referenzen je Zitationsfunktion darstellt und Abbildung 4.2, die die absoluten Häufigkeiten der epistemischen Funktionen pro Kapitel darstellt.

Deutlich erkennbar ist, dass sowohl anwendungsbezogene als auch theoretische Bezüge nahezu ausschließlich in Einleitung und Diskussion, bzw. Fazit vorkommen (siehe 4.2). Das Forschungsproblem betreffende Zitate finden sich zum überwiegenden Teil in der Einleitung, sind aber auch bei der Präsentation der Ergebnisse relevant sowie in der abschließenden Reflektion (Diskussion und Fazit). Zitate, die zum Verweis auf den Forschungsgegenstand oder zur Beschreibung von Eigenschaften empirischer Objekte dienen, werden in allen Kapiteln verwendet, etwas häufiger allerdings in den Einleitungs- und Ergebnisteilen. Ein Großteil der methodischen Verweise wird in Methodenkapiteln getätigt. Daneben sind noch einige empirische Objekt-Bezüge in den Methodenkapiteln enthalten.

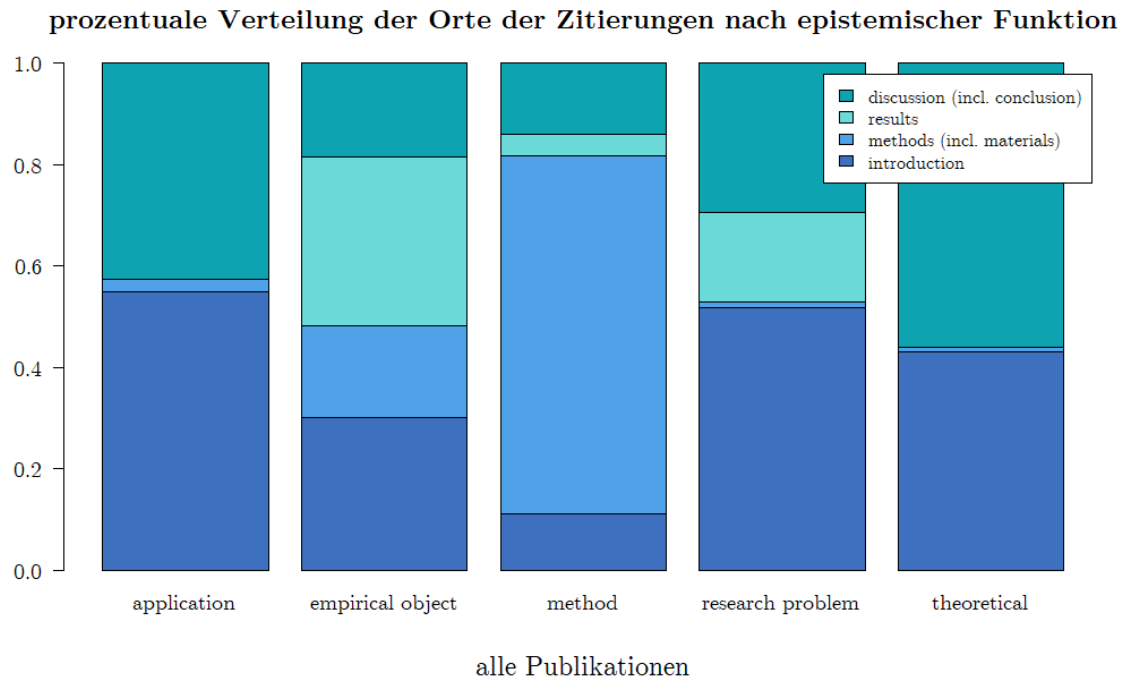


Abbildung 4.1.: Prozentuale Verteilung der Orte der Zitierungen nach epistemischer Funktion

⁴Aufgrund der geringen Besetzung einiger Zellen, kann dieser Zusammenhang nicht mittels Chi²-Unabhängigkeitstest und der Berechnung eines Zusammenhangsmaßes wie Cramer's V überprüft werden. Die Interpretation stützt sich deshalb auf den Vergleich der absoluten und relativen Häufigkeitsverteilungen (Zeilenprozent).

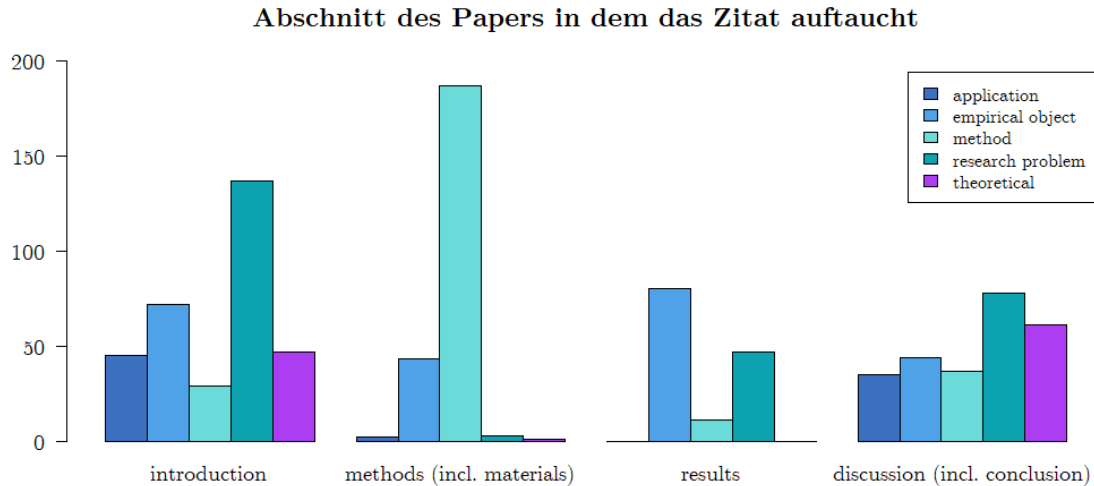


Abbildung 4.2.: Verteilung der Zitate nach standardisierten Kapiteln

Wirft man einen Blick auf die Verteilung der Funktionsbeziehungen bezüglich der Orte ihres Vorkommens nach Dokumententyp getrennt, zeigen sich einige Unterschiede. Zum Beispiel lässt sich die hohe Anzahl methodischer Verweise in den Methodenkapiteln hauptsächlich auf empirische und Modellierungsartikel zurückführen. Die Zitate im Abschnitt „results“ stammen fast ausschließlich (zu 92%) aus dem empirischen Artikel „empirisch_1_2273“.

Diese Unterschiede relativieren sich, betrachtet man die relativen und nicht die absoluten Anteile der Zitationen pro Kapitel in den jeweiligen Dokumentengruppen (siehe Abbildung 4.3). Für alle drei Gruppen ist gleichermaßen beobachtbar, dass anwendungsbezogene, das Forschungsproblem betreffende und theoretische Bezüge überwiegend in Einleitung und Diskussion zu finden sind. Auch die relativen Häufigkeiten unterscheiden sich nur gering und in nicht-aussagekräftiger Art und Weise. Außerdem zeigt sich deutlich, dass methodische Referenzen vorrangig in den Methodenkapiteln vorkommen. Bezüge auf das empirische Objekt finden sich mit Ausnahme der Ergebniskapitel in allen Abschnitten. Dieses Muster zeigt sich homogen und relativ unabhängig von den epistemischen Teilgemeinschaften innerhalb des Feldes der Invasionsbiologie.

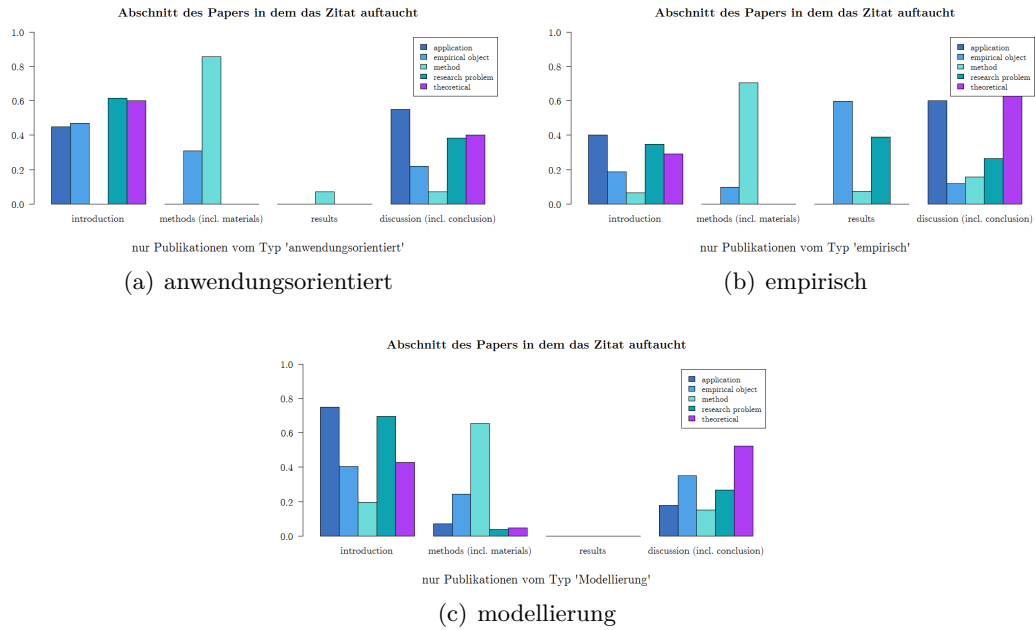


Abbildung 4.3.: Prozentuale Verteilung der Zitate nach standardisierten Kapiteln für die drei Dokumentgruppen

4.1.4. Selbst-Zitation

Von den 907 identifizierten und kodierten Segmenten beziehen sich 101 (rund 11%) auf Quellen, in denen einer der Autoren auch an der Erstellung der jeweils vorliegenden Publikation mitgewirkt hat (siehe Tabelle 4.4). Die Selbstzitationsrate unterscheidet sich leicht zwischen den drei feldinternen Dokumentgruppen, wobei die empirischen Publikationen mit 7% die geringste, die anwendungsorientierten mit 12% eine mittlere und die Publikationen vom Typ „Modellierung“ mit 17% die höchste Rate aufweisen. Betrachtet man die Quoten auf Ebene der untersuchten Artikel fällt aber auf, dass einzelne Publikationen die Zahlen stark beeinflussen. So weichen die Artikel *anwendungsorientiert_1_2159* und *modellierung_2_73* mit 27 und 23 Selbstzitationen (20% und 22%) von der generellen Tendenz ab. Eine Analyse auf Ebene der Dokumentgruppen scheint daher nicht unbedingt aussagekräftig, die generelle Tendenz von rund 11% Selbstzitationsrate scheint jedoch plausibel.

Dokumentname, bzw. Dokumentgruppe	no self-citation	self-citation	summe
anwendungsorientiert_1_2159	108	27	31
anwendungsorientiert_3_573	61	3	14
anwendungsorientiert_4_4261	53	1	12
anwendungsorientiert	222	31	253
empirisch_1_2273	166	11	0
empirisch_2_2845	118	9	48
empirisch_3_2754	99	9	42
empirisch	383	29	412
modellierung_1_1509	67	9	40
modellierung_2_73	81	23	33
modellierung_3_4256	53	9	16
modellierung	201	41	242
gesamt	806	101	907

Tabelle 4.4.: Übersicht absolute Häufigkeiten: self_citation

4.1.5. Zuordnung Felddatensatz und Clusterzurodnung

In den neun ausgewählten und analysierten Publikationen wurden insgesamt 568 verschiedene Referenzen zitiert (siehe Tabelle 4.5). Zu 390 dieser zitierten Referenzen (69%) konnten über das Abgleichen von WoS-IDs weitere Informationen ergänzt werden. Von diesen 390 mit WoS-IDs versehenen Quellen sind 354 (91%) innerhalb des Zeitfensters des „Felddatensatzes Invasionsbiologie“, der im Rahmen eines Forschungsprojekts am DZHW erhoben wurde. Der Felddatensatz beinhaltet alle zwischen den Jahren 2000 und 2019 erschienenen und im Web of Science indexierten Publikationen mit invasionsbiologischem Bezug. Es kann davon ausgegangen werden, dass nahezu alle invasionsbiologisch relevanten Publikationen in dem Felddatensatz enthalten sind. Die restlichen 36 (9%) können aufgrund ihres Alters (Erscheinungsjahr vor 2000) nicht Teil des Felddatensatzes sein. Von den 354 zitierten, mit WoS-IDs versehenen und im Zeitfenster des Felddatensatzes enthaltenen Publikationen, ist mehr als die Hälfte (198 Quellen, entspricht 56%) im Felddatensatz enthalten. Für Sie ist eine eindeutige Zuordnung zu einem der gebildeten Cluster möglich. Die übrigen 156 Referenzen sind demnach nicht im Felddatensatz enthalten, was darauf hinweist, dass es sich hierbei um Zitationen handelt, die auf Wissensbestände aus anderen Feldern verweisen.

	Referenzen		Segmente	
gesamt	568		959	
davon mit WoS-ID	390	68,66%	619	64,55%
davon ohne WoS-ID	178	31,34%	340	35,45%
davon im Zeitfenster des Felddatensatzes	354	90,77%	571	92,25%
davon nicht im Zeitfenster des Felddatensatzes	36	9,23%	48	7,75%
davon im Felddatensatz enthalten	198	55,93%	344	60,25%
davon nicht im Felddatensatz enthalten	156	44,07%	227	39,75%

Tabelle 4.5.: Übersicht Angaben Felddaten

Möchte man die Eigenschaften der im Felddatensatz enthaltenden Referenzen genauer analysieren, eignet sich wiederum eine Betrachtung auf Ebene der kodierten Segmente (siehe Abbildung 4.5). Von den insgesamt 959 kodierten Segmenten (inklusive Doppelkodierungen) können zu 619 Segmenten Informationen aus dem Felddatensatz zugeordnet werden, wovon sich 571 innerhalb des relevanten Erscheinungszeitraums befinden. Von diesen 571 befinden sich 344 (rund 60%) innerhalb des Felddatensatzes und damit innerhalb eindeutig zuordenbarer Cluster in der Invasionsbiologie. 227 (rund 40%) können nicht im Felddatensatz enthalten. Die nicht im Felddatensatz enthaltenen Zitationen verweisen demnach auf Wissensbestände außerhalb des Feldes.

Die Differenz zu den oben berichteten Anteilen auf Ebene einzelner, eindeutiger Referenzen ist nicht ohne weiteres zu erklären. Es bietet sich die Interpretation an, dass Quellen, die an mehreren Stellen in einer Publikation zitiert werden, eher einem der Cluster zugeordnet werden können, als Zitationen, die nur ein einziges Mal im Text vorkommen. Das könnte bedeuten, dass gut vernetzte, etablierte Quellen häufiger mehr als ein Mal im Text zitiert werden, als etwa Quellen, die schlechter vernetzt sind und thematisch keinem Cluster zugeordnet werden können. Es könnte allerdings auch sein, dass die Beobachtung der differierenden Anteile nach Segment-, bzw. Referenzebene nicht überzufällig ist und sich bei Vergrößerung der Stichprobe relativiert. Die Beobachtung könnte als Hypothese für weitere Forschung verwendet werden.

Von denjenigen Zitationen, die keinem Cluster eindeutig zuordenbar sind, sind knapp die Hälfte methodische Zitationen (48%). Bezüge auf das Forschungsproblem und empirische Objekte entfallen je rund 20% der nicht-zuordenbaren Zitationen. Theoretische und anwendungsbezogene Zitationen stel-

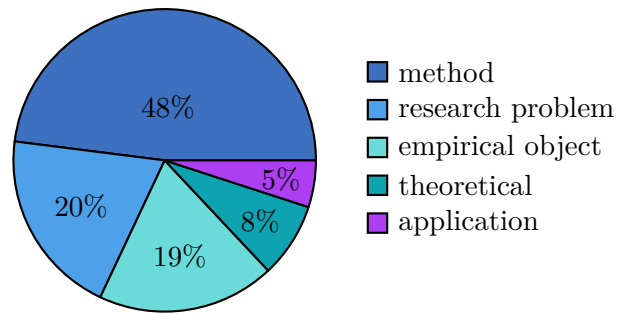


Abbildung 4.4.: Verteilung der keinem Cluster zuordenbaren Zitationen nach epistemischer Funktion in %

len mit 8% (theoretisch) und 5% (anwendungsorientiert) nur einen geringen Teil nicht-zuordenbaren Zitationen dar (siehe Abbildung 4.4).

Theoretische, anwendungsorientierte und das Forschungsproblem betreffende Bezüge sind innerhalb des Felddatensatzes gut vernetzt, sodass sie zu einem überwiegenden Teil thematischen Clustern zugeordnet werden konnten (80% der theoretischen und je 75% der anwendungsorientierten und der das Forschungsproblem betreffenden Quellen). Von den methodischen Referenzen konnten nur etwa ein Drittel (32%) thematischen Clustern zugeordnet werden. Für Referenzen die das empirische Objekt betreffen, ist das Verhältnis ausgewogen (siehe Abbildung 4.5).

Möglichkeit der Zuordnung zu Clustern im Felddatensatz

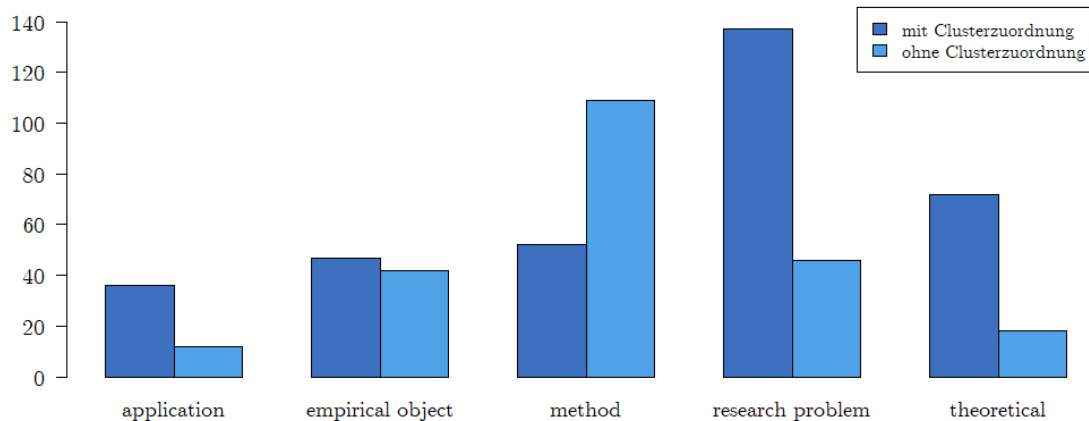


Abbildung 4.5.: Zuordnung zu Clustern im Felddatensatz nach epistemischer Funktion der zitierten Quellen

4.1.6. Erscheinungsjahre der zitierten Quellen

Für die Auswahl des Samples wurden nur Publikationen berücksichtigt, die im Jahr 2019 erschienen sind. Die in den gesammelten Aufsätzen zitierten Referenzen unterscheiden sich selbstverständlich hinsichtlich ihres Erscheinungsjahres. Die älteste zitierte Referenz erschien 1985, die jüngste erschien im

Jahr 2018.⁵ Das arithmetische Mittel liegt bei 2009, der Median, der bei schiefen Verteilungen etwas stabiler gegenüber Ausreißern ist, bei 2010. Die Publikationsjahre der Referenzen bilden eine links-schiefe Verteilung, was bedeutet, dass jüngere gegenüber älteren Publikationen im Sample bevorzugt zitiert werden.

Betrachtet man die Verteilung der Erscheinungsjahre der zitierten Literatur nach zugeordneter epistemischer Funktion, zeigen sich Unterschiede zwischen den Gruppen (siehe Abbildung 4.6). So liegt das mittlere Publikationsjahr von Referenzen mit empirischen Objektbezug bei 2008, mit theoretischen Bezug bei 2010, mit methodischen und anwendungsorientierten Bezug bei 2011 und bei Referenzen, die sich auf das Forschungsproblem beziehen bei 2013. Im Rahmen einer Varianzanalyse wurden die Mittelwertunterschiede auf Signifikanz getestet: die Unterschiede mindestens zweier Gruppen sind höchst signifikant ($p < 0,0001$).⁶ Mittels eines Scheffé-Tests konnte darüber hinaus geprüft werden, auf welche Gruppenunterschiede dieser Zusammenhang zurückzuführen ist. Besonders stark prägt sich dabei der Mittelwertunterschied im Alter der zitierten Literatur von Referenzen mit Bezug auf empirische Objekte gegenüber Referenzen mit Bezug auf das Forschungsproblem aus. Die Mittelwerte unterscheiden sich zwischen den beiden Gruppen um 5 Jahre. Der Unterschied ist höchst signifikant ($p < 0,0001$).

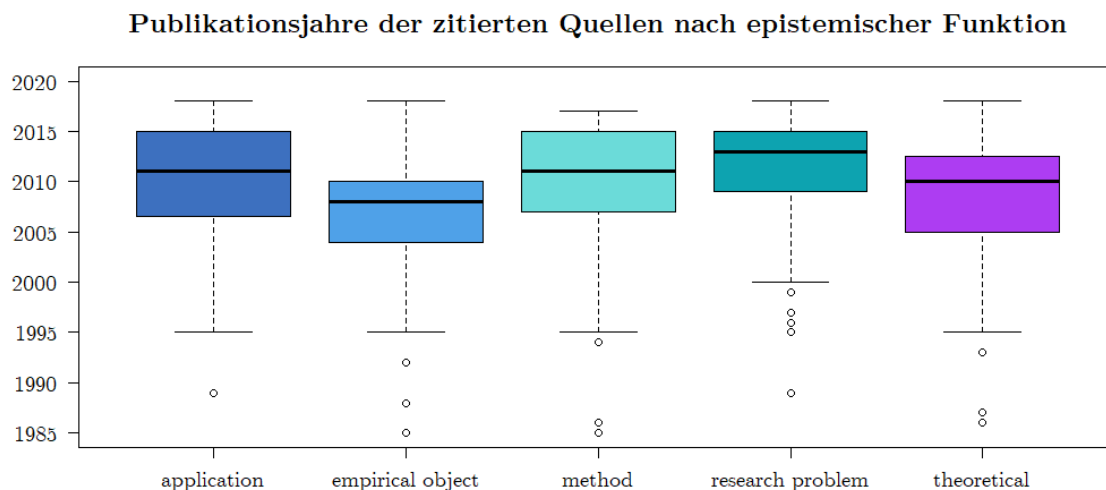


Abbildung 4.6.: Darstellung der Verteilungen von Publikationsjahren nach epistemischer Funktion (Boxplots)

In der Invasionsbiologie ist es üblich bei der Erwähnung oder Beschreibung der Eigenschaften von empirischen Objekten, wie z.B. von bestimmten invasiven Tier- oder Pflanzenarten, diejenige Quelle zu zitieren, in der die Art das erste Mal Erwähnung findet, bzw. das erste Mal ausführlich beschrieben und klassifiziert wird. Diese Tradition konnte auch in den neun untersuchten Artikeln beobachtet

⁵Grundlage für die Auswertung sind hier nur diejenigen Referenzen, die auch im Web of Science verzeichnet sind. Es kann daher zu einer geringen Verzerrung kommen, da davon auszugehen ist, dass ältere Publikationen eine etwas schlechtere Abdeckung in der Datenbank haben als jüngere Veröffentlichungen. Diese Verzerrung muss bei der Interpretation der Ergebnisse und der Bewertung ihrer Aussagekraft berücksichtigt werden.

⁶Der Test setzt Normalverteilung der metrisch skalierten abhängigen Variable und Varianzhomogenität innerhalb der Gruppen voraus. Die Publikationsjahre der Referenzen sind, wie oben beschrieben, nicht eindeutig normalverteilt. Aufgrund der hohen Fallzahl, sollte der Test dennoch robust gegen Ausreißer sein. Zur Überprüfung der Annahme auf Varianzhomogenität wurde der Levene Test durchgeführt. Mit $p = 0,1777$ ($F = 1,581$ und $Df = 4$) muss die Nullhypothese, dass Varianzhomogenität vorliegt, angenommen werden.)

werden. Sie könnte dafür verantwortlich sein, dass Zitate, die sich explizit auf empirische Objekte beziehen, durchschnittlich vergleichsweise älter sind, als z.B. Zitate, die sich dem Forschungsproblem widmen. Aufgrund der großen Spannweite und einigen Ausreißern nach unten, könnte es sogar sein, dass das durchschnittliche Alter empirischer Objektbezüge unter Berücksichtigung der nicht in Web of Science indexierter Publikationen, noch etwas nach unten korrigiert werden müsste. Dahingegen lässt die deutlich geringere Spannweite der auf das Forschungsproblem bezogenen Referenzen vermuten, dass sich die bislang nicht in Web of Science befindlichen Publikationen hier nicht so stark auswirken würden. Verweise mit Bezug auf das Forschungsproblem sind durchschnittlich deutlich jünger. Die zum größeren Teil in den Einleitungskapiteln angeführten Quellen dienen z.B. der Wiedergabe des Forschungsstandes. Dass sich dabei eher auf aktuellere, jüngere Artikel gestützt wird, als auf tradierte, könnte z.B. darauf zurückgeführt werden, dass in älterer Literatur keine relevanten Beiträge zu den aktuellen Forschungsproblemen enthalten sind. Bei den aktuell verhandelten Forschungsproblemen könnte es sich um so neue und innovative Fragestellung handeln, dass diese in älteren Quellen noch gar nicht thematisiert wurden. Es könnte allerdings auch der Fall sein, dass zwar relevante ältere Literatur vorhanden ist, diese aber zugunsten der Zitation jüngerer Literatur vernachlässigt wird. Dies könnte dafür sprechen, dass die Herstellung von Aktualität und damit verbunden sicher auch die Begründung von Relevanz, ein der vollständigen Wiedergabe der relevanten Literatur übergeordnetes Ziel ist. Hier wäre sicher ein Feldvergleich zu etwa geistes- und sozialwissenschaftlichen Fachgemeinschaften vielversprechend und aufschlussreich.

4.2. Analyse der Netzwerke

4.2.1. Überblick

Für die Auswertung wurde ausgehend vom Seed Dokument zunächst ein Primärnetzwerk erstellt, das alle zitierten und zitierenden Publikationen (als Knoten) und deren Verbindungen untereinander (als Kanten) enthält. Für eine genauere Betrachtung wurde das Primärnetzwerk anschließend in Teilnetzwerke gegliedert, die jeweils alle Artikel, die über eine bestimmte epistemische Funktion miteinander verbunden sind, enthalten.

Das Primärnetzwerk enthält das Seed Dokument (SD), alle 12 Publikationen, die im Seed Dokument zitiert werden (gekennzeichnet mit r1–r12) und alle 22 Publikationen, die das Seed Dokument zitieren (gekennzeichnet mit z1–z22). Insgesamt enthält das Primärnetzwerk damit 35 Knoten (1+12+22). Innerhalb der 35 Publikationen werden 282 Verbindungen in Form von Zitationen verwirklicht. Da oftmals in einem Artikel gleich mehrfach auf den selben anderen Artikel verwiesen wird, ergeben sich nur 140 Verbindungen mit differierenden Start- und Zielknoten. In der visuellen Darstellung des Netzwerks werden somit 140 verschiedene Kanten realisiert. Die Stärke der Kanten spiegelt die summierten Verbindungshäufigkeiten wider.

Von den ursprünglich 282 Zitationen, die auf die Publikationen des Primärnetzwerks entfallen, betreffen 179 das Forschungsproblem, 79 Zitationen drücken eine Verbindung über empirischen Objekte aus, 13 Verbindungen sind auf theoretische Inhalte und elf methodische Verweise bezogen. Innerhalb der 35 Publikationen werden keinerlei anwendungsbezogene Zitationen getätigt.

Die Abbildung 4.7 zeigt das gesamte Primärnetzwerk in grau und jeweils die Teilnetzwerke der vier verschiedenen epistemischen Funktionen farbig hervorgehoben. In allen Darstellungen ist die Kantenstärke als Summe der Einzelverbindungen proportional dargestellt. Die Größe der Knoten entspricht dem Grad, d.h. der Anzahl eingehender plus ausgehender Verbindungen des jeweiligen Knotens. Für

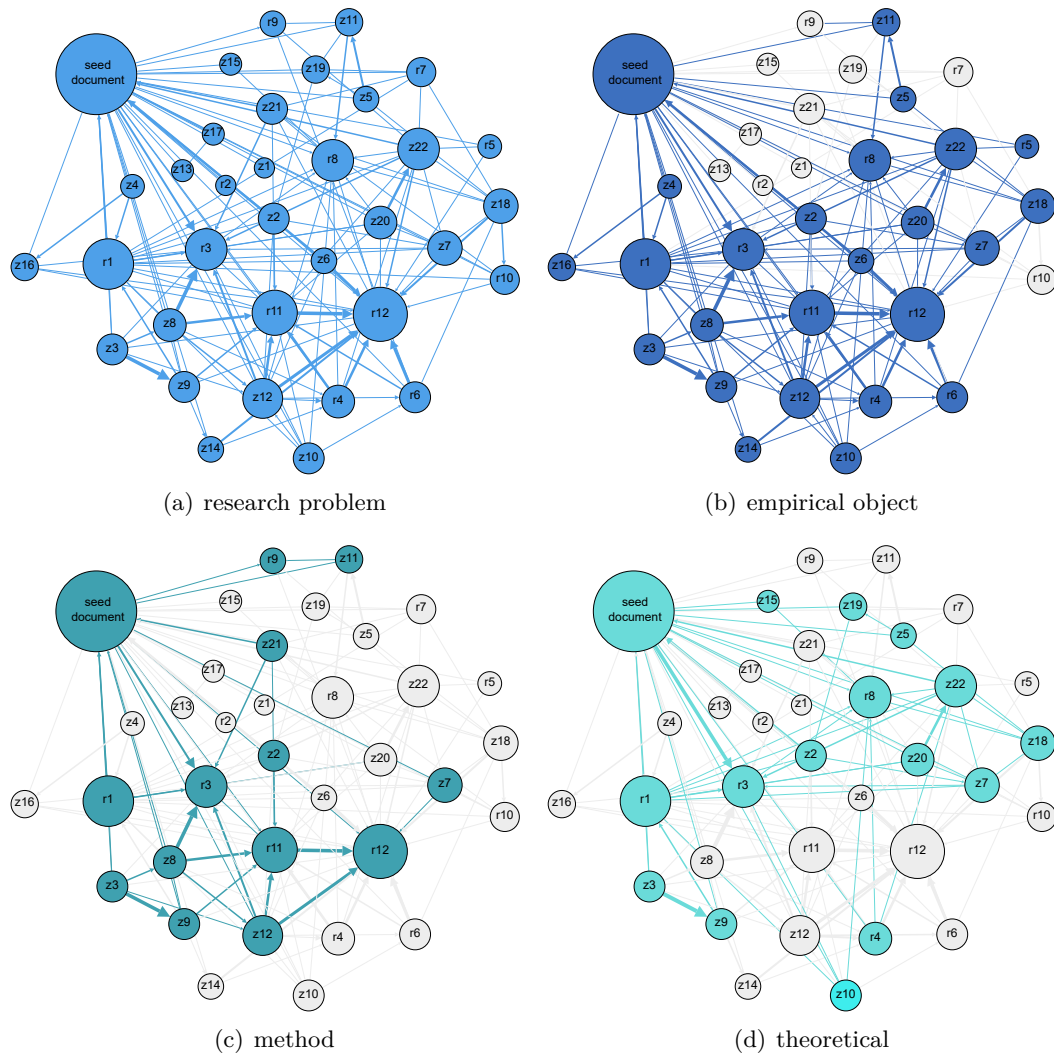


Abbildung 4.7.: Primärnetzwerk mit Hervorhebung der in den Teilnetzwerken enthaltenen Publikationen und Verbindungen

eine übersichtliche Anordnung des Netzwerks wurde in Gephi (Bastian, Heymann und Jacomy, 2009) die integrierte Layout-Funktion „Force Atlas 2“ (Jacomy u. a., 2014) in Kombination mit der Layout-Funktion „Noverlap“ angewendet. Anhand der Darstellungen ist die Größe der jeweiligen Teilnetzwerke deutlich zu erkennen. So sind alle im Primärnetzwerk vorhandenen Publikationen mindestens einmal über die Zitation eines Forschungsproblems mit mindestens einer weiteren der 35 Publikationen verbunden. Alle bis auf drei Publikationen (z9, z10, z15) sind direkt mit dem Seed Dokument über die Zitation eines Forschungsproblems verbunden. Bei Zitierungen, die sich auf empirische Objekte beziehen, sind 25 der 35 Publikationen involviert. Methodische Beziehungen bestehen zwischen 14 der 35 Dokumente und theoretische Verweise werden zwischen 16 der 35 Artikel hergestellt. Die Abbildung verdeutlicht außerdem, dass viele der Knoten aufgrund gleich mehrerer epistemischer Funktionen mit anderen Publikationen verbunden sind. So tauchen etwa die Artikel r1, r3, z2, z3, z7, z9 in allen vier Netzwerken auf. Die Artikel r2, r7, r10, z1, z13, z17 sind dahingegen allein durch Verweise auf Forschungsprobleme im Primärnetzwerk verbunden.

Die Abbildung 4.8 zeigt die Verteilung der Anteile der eingehenden Verbindungen nach epistemischer Funktion gegliedert. Im Gegensatz zu den in Abbildung 4.7 hervorgehobenen Knoten sind hier nur diejenigen Publikationen farblich hervorgehoben, die in einer der anderen Arbeiten zitiert werden.

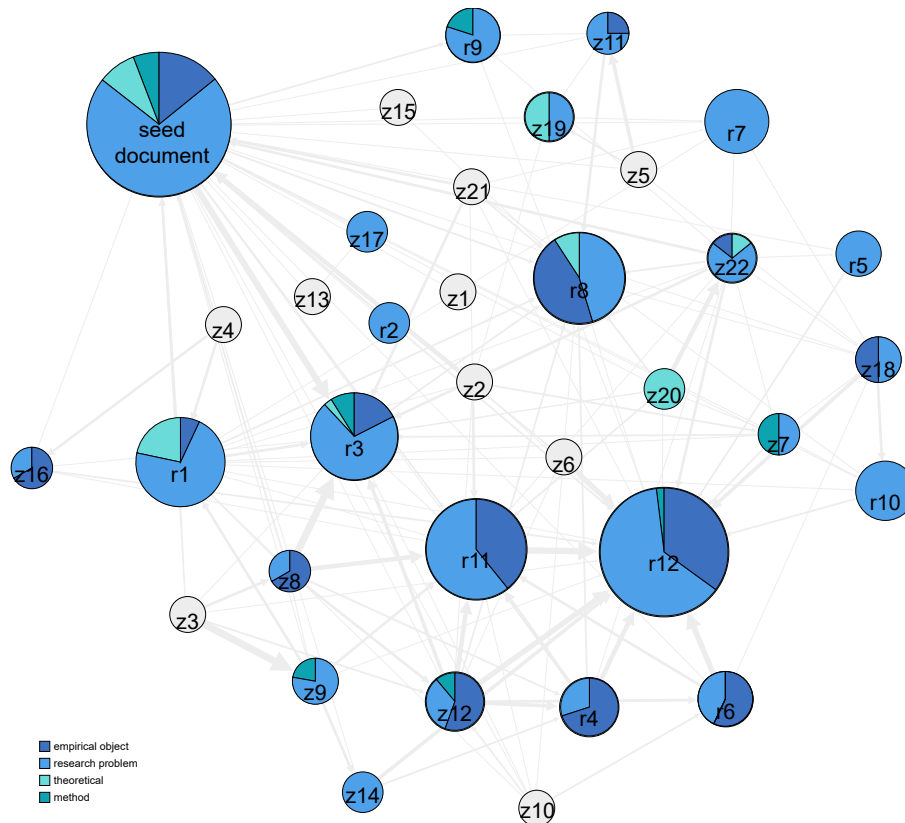


Abbildung 4.8.: Primärnetzwerk mit Hervorhebung der Anteile eingehender Verbindungen nach epistemischen Funktion

Daran lässt sich erkennen, welche Arbeiten besonders für Inhalte welcher epistemischen Funktionen relevant sind. Es lassen sich so einige Publikationen identifizieren, die ausschließlich für Inhalte zu Forschungsproblemen zitiert werden (r2, r5, r7, r10, z14, z17). Ein Aufsatz wird ausschließlich für theoretische Inhalte zitiert (z20). Insgesamt zeigt sich aber, dass insbesondere Publikationen, die mehrfach zitiert werden (dargestellt durch verhältnismäßig größere Kreise) zumeist für verschiedene Inhalte zitiert werden. Mit der Anzahl eingehender Zitierungen steigt somit auch die Diversität unter den verwendeten Funktionen (siehe insbesondere r1, r3, r8, r11, r12).

Das Seed Dokument beschäftigt sich thematisch mit den Auswirkungen von Invasion, genauer, mit den Folgen, die auf die Verbreitung des Invasors „Dikerogammarus Villosus“ zurückzuführen sind. Dikerogammarus Villosus, der große Höckerflohkrebs, ist eine invasive Art von Flusskrebs, der sich seit den 1990er Jahren in westlichen europäischen Flüssen und Seen verbreitet, darunter der Main-Donau-Kanal, der Rhein, der Main, die Donau, die Elbe, die Oder und der Bodensee. Heimisch ist die besonders anpassungsfähige Krebsart im Ponto-Kasparischen Raum in Osteuropa, in den Flussmündungen zum Schwarzen Meer. Als „Killer Shrimp“ bezeichnet, wird er häufig für die Verdrängung anderer Krebsarten und damit einem Wandel in der Biodiversität verantwortlich gemacht. Beispielsweise kann für den Rhein beobachtet werden, dass sich die Verbreitung und Zusammensetzung der Krebsarten seit der Invasion durch Dikerogammarus Villosus in 1994/1995 deutlich verändert hat. Im Seed Dokument versuchen die Autoren herauszufinden, ob sich diese Beobachtung auf interspezifische Konkurrenz (interspecific competition) und Intraguild Predation, also das gegenseitige Angreifen und Fressen, zurückführen lässt.

Während das Seed Dokument noch eine recht klassisch invasionsbiologische Studie darstellt, sind viele

der im Seed Dokument zitierten und das Seed Dokument zitierenden Artikel eher der Kategorie „invasion aware“ zuzuordnen (zur Klassifizierung feldinterner Forschungsansätze siehe Abschnitt 3.1.1, S. 23). Häufig geht es darum die biologischen und ökologischen Eigenschaften bestimmter invasiver oder heimischer Arten zu beschreiben und zu vergleichen. Die allermeisten der im Primärnetzwerk enthaltenen Publikationen beschäftigen sich mit Amphipoden (Ordnung) oder Gammariden (Unterordnung). Viele untersuchen spezifische Eigenschaften einzelner Arten, deren Verhalten in Gemeinschaft mit anderen Arten oder das Invasionsverhalten und die Verbreitung in neue Invasionsgebiete. Theoretische Zitationen beziehen sich häufig auf allgemeine Mechanismen von Invasionsprozessen, die den Status theoretischer Grundannahmen erreicht haben. Methodische Zitationen werden häufig angeführt um auf Details zur Sampling-Strategie hinzuweisen, um das eigene methodische Vorgehen oder das Vorgehen anderer zu kritisieren oder um auf die Nachnutzung bestimmter Daten, Verfahren oder Parameter hinzuweisen. Zitationen, die sich auf empirische Objekte beziehen, werden häufig angeführt um Aussagen zu Eigenschaften bestimmter Arten, Gattungen, Familien und dergleichen zu belegen, bzw. auf frühere Forschung zu den Tieren und Sachverhalten hinzuweisen. Das Forschungsproblem betreffende Zitationen verweisen zumeist auf vorhergehende Forschungen, die sich mit einem ähnlichen oder demselben Forschungsproblem auseinandergesetzt haben.

Um die in den Artikeln bearbeiteten Themen abzubilden, eignet sich die Betrachtung je einzeln nach epistemischer Funktion der Zitierungen. Es wurden Teilnetzwerke erstellt, die jeweils nur die Publikationen und Zitationen enthalten, die einer der vier Funktionen zugeordnet sind (theoretical, method, empirical object und research problem). Die Themen, die sich durch eine genauere Betrachtung des Kontextes von Zitationen ermitteln ließen, sind als Kantenbeschriftungen in die grafischen Darstellungen integriert.

4.2.2. Epistemische Funktion: theoretical

Das Teilnetzwerk für die theoretischen Zitierungen besteht aus 16 Knoten, die über zwölf Kanten miteinander verbunden sind. Um den Graph visuell zu strukturieren und eine erste Idee über mögliche Cluster zu erhalten, wurden die Knoten entsprechend ihrer Modularity Class eingefärbt. Über die Berechnung der Modularität konnten fünf Cluster ermittelt werden. Drei der Cluster sind jeweils unverbunden (hellgrün, dunkelgrün, lila), zwei Cluster sind über einen gemeinsamen Knoten in der Mitte verbunden (hellblau, dunkelblau, verbunden durch z9).

Die Kantenbeschriftungen geben Aufschluss darüber auf welche Themen sich die theoretischen Zitationen beziehen. So ist deutlich zu erkennen, dass das Thema „competition“ eine übergeordnete Rolle spielt. Zumindest enthalten sieben der zwölf Zitierungen den Begriff „competition“ in der Themenbezeichnung. „Competition“ bezeichnet im Kontext der untersuchten Studien den Wettbewerb zwischen verschiedenen Arten, die sich einen Lebensraum teilen. Wettbewerb zwischen Arten kann sich direkt (gegenseitiges Angreifen oder Fressen, „predation“) oder indirekt (durch den Wettbewerb um geteilte Ressourcen) auf die Überlebensfähigkeit der beteiligten Arten auswirken und eingeschränkte (Veränderung der Verteilungshäufigkeiten, „impact on benthic communities“) bis sehr drastische Auswirkungen (komplette Verdrängung einer Art, „replacement“) haben. Im Netzwerk werden verschiedene Aspekte von Wettbewerb sichtbar. Unter dem Label „competition and impact on benthic communities“ summieren sich Verweise, die noch etwas allgemeiner auf die möglichen Auswirkungen von Invasionsprozessen auf die Struktur und Zusammensetzung von Gemeinschaften von Wasserlebewesen hindeuten. Unter dem Label „competition and replacement“ fassen sich Verweise zusammen, die dahingegen konkreter auf eine bestimmte Folge von Invasion hinweisen; nämlich auf die gänzliche Verdrängung einer bestimmten Art durch eine andere (invasive) Art.

in neue Umgebungen einwandern indem sie als blinde Passagiere auf Güterschiffen und Hobby-Yachten mitfahren. Auch der Ausbau von künstlich angelegten Kanälen zählt als eine der Hauptursachen für die Invasion durch nicht-heimische Arten. Die Distribution invasiver Arten nachzuvollziehen und ihre Ursachen zu erklären ist ein Kernthema invasionsbiologischer Forschung mit breiter theoretischer Grundlage.

Das Seed Dokument ist über eingehende Zitationen mit drei anderen Publikationen verbunden, ausgehende Verbindungen bestehen nicht. Auffällig ist, dass alle Zitationen, die auf das Seed Dokument verweisen, sich auf das selbe Topic beziehen: „competition and impact on benthic communities“. Es ist durchaus nicht selbstverständlich, dass eine Publikationen stets für ein und denselben „knowledge claim“ zitiert wird. Im vorliegenden Fall scheint es allerdings so zu sein, dass das Seed Dokument einen recht eindeutigen theoretischen Fokus bereitstellt, für den es von anderen Autoren zitiert wird.

4.2.3. Epistemische Funktion: method

Das Teilnetzwerk für die methodischen Zitierungen besteht aus 14 Knoten, die über neun Kanten miteinander verbunden sind. Die verschiedenen Farben repräsentieren die über die Modularität in Gephi berechneten Cluster (siehe Abbildung 4.10). Es konnten auf diese Weise sechs Cluster berechnet werden. Drei der Cluster (lila, gelb, hellblau) bestehen jeweils lediglich aus zwei Knoten, die durch eine einfache Kante miteinander verbunden sind. Ein weiteres Cluster (dunkelblau) besteht aus drei Knoten, wobei der mittlere (SD) über zwei eingehende Verbindungen mit den anderen beiden verbunden (r1, z21) ist. Das Größte der zusammenhängenden Teile des Netzwerks besteht aus fünf Knoten, die nach Modularität zwei Clustern zugeordnet werden. Die Publikation r3 verbindet die beiden Cluster (dunkel- und hellgrün) in der Mitte. Zwei Kanten sind etwas dicker dargestellt (z3–z9 und z8–r3), da hierbei die Quellen gleich mehrfach, jeweils zweimal, im Startknoten zitiert werden.

Inhaltlich thematisieren die Zitationen allgemein den methodischen Zugang („methodical focus in previous studies“) oder konkret die Verwendung von Methoden („SIA method“) und Modellparametern („model calculations“, „methodical parameters for experiments“). Methodenbezogene Zitationen enthalten durchaus auch kritische Auseinandersetzungen („Methodenkritik“). Dabei ist im vorliegenden Fall die zweifache Verwendung des Topics „Methodenkritik“ vielleicht etwas irreführend: in z3 wird z9 als Beispiel für methodische Studien angeführt, die sich aus einem qualitativen Ansatz, mit geringer Stichprobengröße einem Forschungsproblem nähert. Vor der Folie dieser (und weiterer) Studien soll insbesondere der Verdienst von quantitativen Ansätzen hervorgehoben werden, wie es beispielsweise durch r3 und die vorliegende Arbeit (z3) verwirklicht wird. Dabei ist interessant zu wissen, dass z3 und r3 einen gemeinsamen Autor teilen (René Gergs). Das kritische Element ist sicher in der Zitation enthalten, durch die genauere Analyse der Zusammenhänge und inhaltlichen Verknüpfung würde sich aber ggf. auch eine etwas anders zugeschnittene Benennung der Topics anbieten. Durch etwa „methodischer Ansatz: quantitativ“ und „methodischer Ansatz: qualitativ“ könnten hier Unterschiede in der methodischen Herangehensweise differenziert werden. Darüber hinaus würden sich ggf. auch noch andere Topics in die Unterscheidung quantitativ/qualitativ (als Oberkategorie) überführen lassen.

Deutlich von den konkret auf Methoden, methodische Modellparameter und Berechnungsmodelle abzielenden Zitationen lassen sich Verweise abgrenzen, die sich auf das Sampling beziehen. Hierbei ist auffällig, dass es sich bei allen auf das Sampling beziehenden Zitationen um solche handelt, bei denen mindestens ein Autor der Autorenlisten gleich ist. In r11 (Dick, Platvoet und Kelly, 2002) wird r12 (Dick und Platvoet, 2000) zitiert um auf genauere Beschreibungen der Orte der Datenerhebung hinzuweisen. Die Publikationen r1 und z21 verweisen auf das Seed Dokument um auf die Nachnutzung

früher erhobener Daten hinzuweisen und betreffen damit ebenfalls die Datenerhebung. Das Seed Dokument und die Artikel r1 und r21 teilen zwei Autoren einer niederländischen Arbeitsgruppe: Gerard van der Velde und Abraham bij de Vaate.

Das Seed Dokument ist über zwei eingehende Verbindungen mit anderen Publikationen verbunden. Beide methodischen Zitierungen beziehen sich dabei auf die Datenerhebung. Einmal als konkrete Nachnutzung der Daten und einmal als Verweis auf ergänzende Informationen zum Sampling-Gebiet. Bei beiden Zitierungen handelt es sich um Selbstzitationen, was die Bedeutung des Einflusses des Seed Dokuments hinsichtlich methodischer Aspekte schmälert.

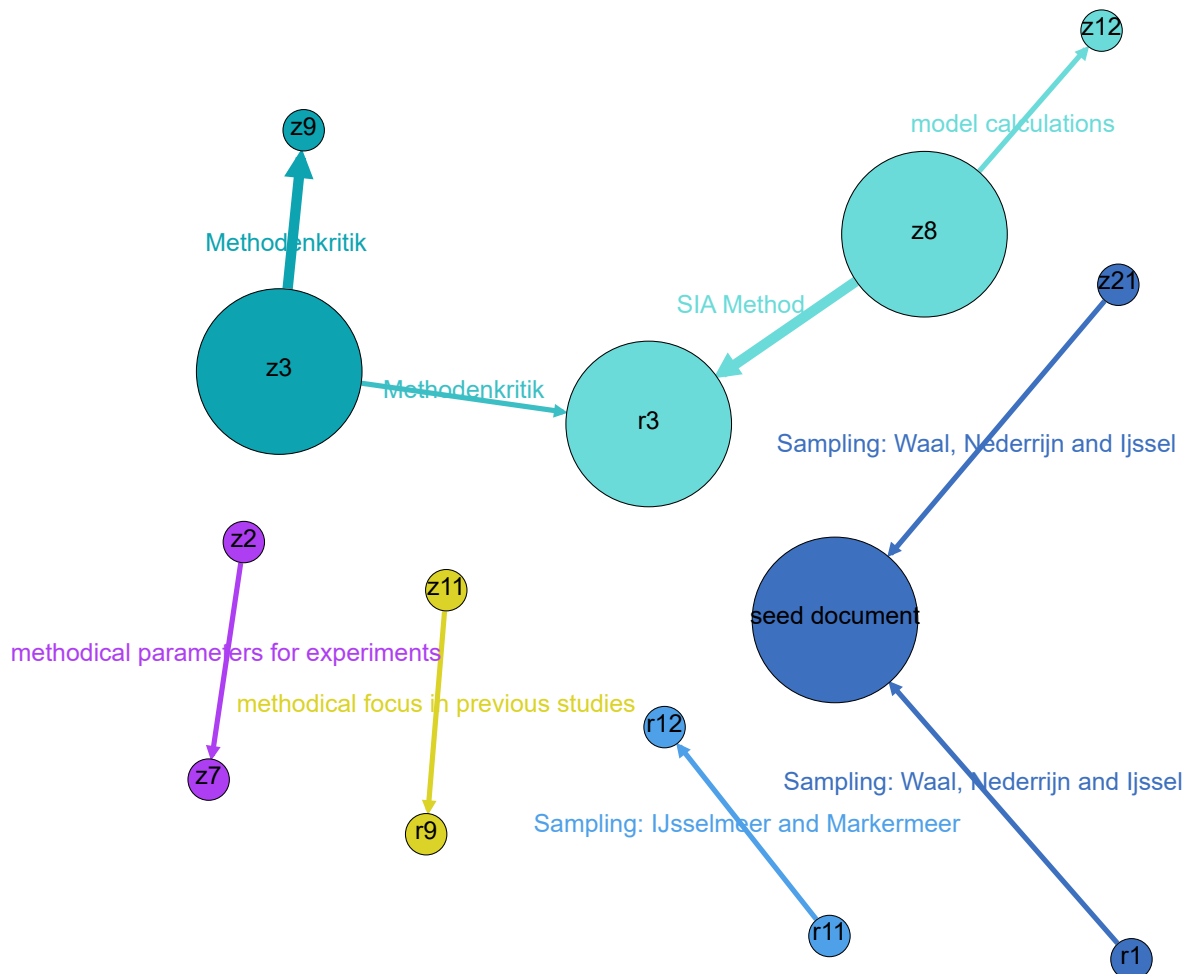


Abbildung 4.10.: Teilnetzwerk „method“

Insgesamt zeigt sich, dass sich hinsichtlich methodischer Zitierungen ein nur lose verknüpftes Netzwerk mit recht diversen Topics beschreiben lässt. Die Heterogenität in der Benennung der Topics könnte durch ein höheres Abstraktionsniveau noch verdichtet werden. So könnten etwa quantitative und qualitativen methodischen Ansätzen unterschieden werden. Von diesen wiederum wären Zitationen, die sich mit Datenerhebung beschäftigen abzugrenzen. Es ist zu vermuten, dass bei einer Erweiterung der Datengrundlage sich noch etwas klarere Strukturen im Netzwerk etablieren könnten. Etwa, dass Verweise auf konkrete Methoden, wie der SIA Methode, eigene Cluster bilden. Auch wäre bei einer größeren Datengrundlage interessant zu beobachten, ob sich ortsbezogene Cluster hinsichtlich der Orte der Datenerhebung identifizieren lassen. Der vorliegende Fall deutet die Entstehung eines „niederländischen Clusters“ schon an.

4.2.4. Epistemische Funktion: empirical object

Für Zitationen, die sich auf das epistemische Objekt beziehen wurden zwei verschiedene Ansätze für die Ermittlung der Topics verfolgt (siehe Abbildungen 4.11 und 4.12). Zitationen zu empirischen Objekten werden zumeist angeführt um die Eigenschaften bestimmter Arten zu belegen, bzw. um auf frühere Forschungsergebnisse zu den Eigenschaften der Arten hinzuweisen. Daraus ergibt sich für thematisches Mapping ein doppelter Informationsgehalt: Zum einen ist die Information relevant um welche empirischen Objekte es sich denn eigentlich handelt, welche empirischen Objekte Gegenstand der Zitation sind. Und zum anderen ist die Information relevant, welche Eigenschaften thematisiert werden bzw. welche weitergehenden Informationen zu den empirischen Objekten thematisiert werden. Aus diesem doppelten Informationsgehalt bezüglich auf empirische Objekte bezogene Zitationen ergibt sich die Chance einer getrennten Auswertung der beiden Dimensionen.

Zunächst aber zur generellen Beschreibung des Netzwerks: Das Teilnetzwerk für die Zitationen, die sich auf das empirische Objekt beziehen, besteht aus 25 Knoten, die über 50 Kanten miteinander verbunden sind. Über die Berechnung der Modularität konnten vier Cluster ermittelt werden, die jeweils farblich hervorgehoben sind (siehe Abbildungen 4.11 und 4.12). Die vier Cluster unterscheiden sich in ihrer Größe nur gering. Das größte der vier Cluster (hellblau) enthält acht Publikationen, zwei Cluster (dunkelblau, lila) enthalten je sechs Publikationen und das kleinste Cluster (gelb) enthält fünf Publikationen. Insgesamt sind die blauen Cluster deutlich stärker vernetzt als das lila und gelbe Cluster, die sich jeweils eher am Rand des Netzwerkes ansiedeln lassen und jeweils nur zwei Knoten haben, die in das Netzwerk hinein verweisen und so mit den anderen Clustern verbunden sind. Die Darstellung des Netzwerks unterscheidet sich nicht zwischen den beiden Topic-Ansätzen, da jeweils nur die Kantenbeschriftungen verschieden definiert wurden. Dies begünstigt den Vergleich der beiden Ansätze.

Abbildung 4.11 zeigt die aus den Kontexten der Zitationen ermittelten Bezüge auf konkrete empirische Objekte. Die Kanten sind zumeist mit den lateinischen Namen der thematisierten Art beschriftet. In einigen Fällen wurde auf mehr als eine Art referiert, sodass sich auch mehrere Arten in den Beschriftungen wiederfinden. Der Invasor *D. Villosus* wurde beispielsweise einige Male gemeinsam mit *D. Haemobaphes* erwähnt, einer Art der selben Gattung und vielen ähnlichen Eigenschaften. An zwei Stellen wurde auf Gammariden allgemein Bezug genommen (z20–r8⁷, z4–z16). An zwei Stellen wurde der Fluss Rhein als empirisches Objekt thematisiert (z8–r1, z4–r1). Die allermeisten der Zitationen beziehen sich auf den Invasor *D. Villosus*, der auch das Untersuchungsobjekt des Seed Dokuments ist. Während sich die Zitationen zu *D. Villosus* hauptsächlich in den beiden blauen Clustern finden, sind andere Arten wie z.B. *G. Pulex*, *G. Fossarum* und *G. Roeseli* (alles Arten aus der Familie der Gammariden) fast ausschließlich im grünen Cluster zu finden. Das gelbe Cluster weist zweimal das Topic „the river rhine“ auf, was darauf hindeutet, dass hier der Fokus der Zitationen etwas anders gelagert ist. Bei diesen Zitationen wird der Lebensraum zum Untersuchungsobjekt, unabhängig von einer bestimmten Art, die diesen besiedelt.

Die Abbildung 4.12 zeigt die Eigenschaften und sonstige Informationen zu den empirischen Objekten, die in den Zitationen thematisiert werden. Am häufigsten sind die Kategorien, die eine Art als Raubtier oder Allesfresser beschreiben („predator“, „omnivorous predator“). Daneben werden viele Zitationen angeführt um die Ausbreitung von Arten zu belegen („distribution“). Darüber hinaus werden auch konkrete Eigenschaften von Arten thematisiert, wie etwa deren bevorzugter Lebensraum („habitat preference“) oder ihre Anpassungsfähigkeit („adaptability“). Auch in dieser Abbildung zeigen sich

⁷Die hier und im Folgenden gewählte Schreibweise ist wie folgt zu lesen: von Startknoten zu Zielknoten. D.h. Von z20 zu r8 und von z4 zu z16.)

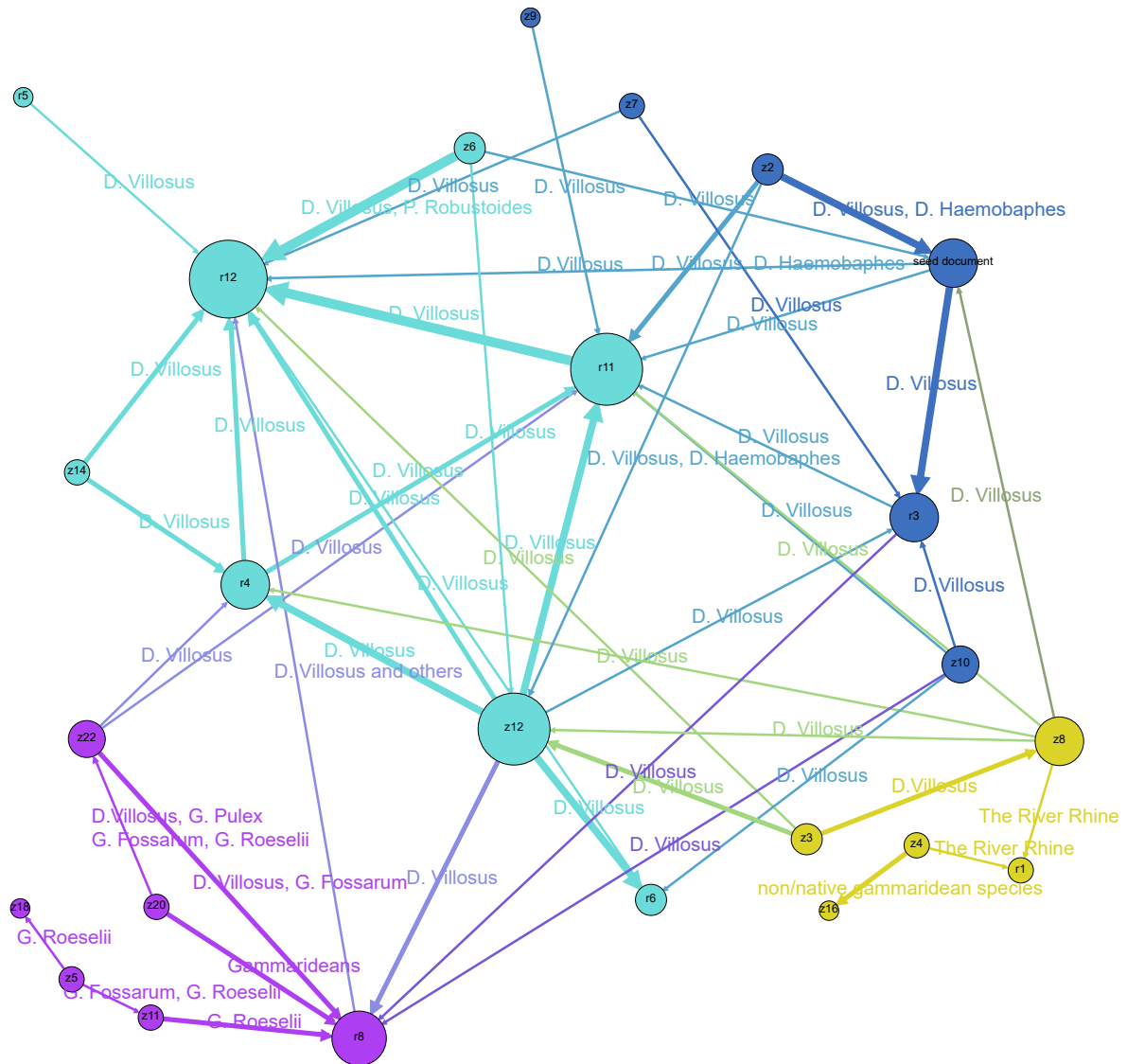


Abbildung 4.11.: Teilnetzwerk „empirical object“, Topics 1: Name

die beiden blauen Cluster als inhaltlich nahe beieinanderliegend. Die häufigsten Themen beziehen sich auf die Eigenschaft des Raubtier-Seins („predator“, „omnivorous predation“) und auf dessen negative Auswirkung auf das Ökosystem („replacement“). Von den beiden blauen Clustern heben sich inhaltlich sowohl das gelbe als auch das lila Cluster ab. Innerhalb des gelben Clusters finden sich einige der Zitationen, die die Verteilung von Arten („distribution“) thematisieren. Das lila Cluster ist gekennzeichnet durch die Topics „adaptability“ und „habitat preference“.

Das Seed Dokument zitiert drei andere Dokumente für Eigenschaften von *D. Villosus*. Zweimal wird dabei dessen Vorliebe für bestimmte Umgebungen thematisiert und einmal steht dessen Klassifikation als Raubtier im Vordergrund. Das Seed Dokument wird von drei anderen Artikeln als Quelle für Informationen zu empirischen Objekten angeführt. Dabei geht es ebenfalls zweimal um den bevorzugten Lebensraum des *D. Villosus* (davon einmal auch um *D. Haemobaphes*) und einmal um *D. Villosus* als allesfressender Räuber.

Führt man die beiden Abbildungen interpretativ zusammen, zeigen sich die blauen Cluster geprägt von Zitationen, die *D. Villosus* als räuberischen Allesfresser beschreiben. Wenn es dahingegen um andere Arten von Gammariden geht, wie sie im lila Cluster vermehrt auftreten, scheinen Eigenschaften



4.2.5. Epistemische Funktion: research problem

52

Abbildung 4.13). Insgesamt können fünf Cluster unterschieden werden. Drei der Cluster sind mit jeweils drei Knoten recht klein (gelb, lila, dunkelgrün), die zwei übrigen Cluster mit zehn (hellblau) und 17 Knoten (dunkelblau) verhältnismäßig größer.

Der Abbildung 4.13 liegen insgesamt 179 Zitationsverbindungen zu Grunde. In der Grafik visualisiert sind davon nur 110, da bei Zitierungen mit gleichem Start- und Zielknoten die Kantenstärke entsprechend gewichtet wurde. Betrachtet man allerdings die ermittelten Topics der gesamten 179 Zitationen, lassen sich die wichtigsten, d.h. in diesem Falle häufigsten, Themen herausarbeiten. Etwa 80% der Kantenbeschriftungen verteilen sich auf die neun Topics: „replacement“, „IGP“ (intra-guild predation), „environmental factors“, „distribution“, „habitat preference“, „invasion success“, „change and replacement“, „impact on benthic community“ und „impact on ecosystem“ (siehe Tabelle 4.6).

Topic	absolute Häufigkeit	relative Häufigkeit
„replacement“	38	21%
„IGP, intraguild predation“	21	12%
„environmental factors“	19	11%
„distribution“	16	9%
„habitat preference“	14	8%
„invasion success“	12	7%
„change and replacement“	11	6%
„impact on benthic community“	7	4%
„impact on ecosystem“	7	4%
andere	34	18%
Summe	179	100%

Tabelle 4.6.: Übersicht Topics „research problem“ (häufigste 82%)

Thematisch liegen die Topics recht nah beieinander und sind nicht immer vollständig trennscharf. Die bei der Analyse berücksichtigten Publikationen beschäftigen sich inhaltlich zum größten Teil mit den Voraussetzungen und Folgen von Invasion durch Gammariden. Dabei gilt ein besonderer Fokus der Invasion mitteleuropäischer Gewässer durch den Invasor *Dikerogammarus Villosus* und dem Rückgang heimischer oder etablierter Gammariden-Arten. Häufig werden Veränderungen in Ökosystemen („impact on ecosystem“) und ihrer Bevölkerung mit verschiedenen Arten („impact on benthic community“) auf Invasion zurückgeführt. Es wird betrachtet welche Arten seit wann wo vorkommen („distribution“) und welche Umweltbedingungen („environmental factors“, „habitat preference“) ihre Etablierung („invasion success“) begünstigen. Während einige die Durchsetzungsfähigkeit und weite Verbreitung von Invasoren auf deren aggressives Verhalten gegenüber heimischen und etablierten Arten („IGP, intraguild predation“, „predation“) zurückführen, betonen andere deren Anpassungsfähigkeit („adaptability“) in neuen Umgebungen. Die Verdrängung von Arten („replacement“) durch Invasoren wird in vielen der im Netzwerk berücksichtigten Publikationen thematisiert. In diesem Zusammenhang wird auch häufig die Hierarchie der Nahrungsketten thematisiert („trophic position“, „food web“, „gammarideans as food, trophic function“). Das Zusammentreffen invasiver und heimischer oder etablierter Arten ist aber keinesfalls immer konfliktträchtig. In einigen Fällen wird auch über friedliche Begegnungen und Koexistenz („coexistence“, „interaction“) berichtet.

Die Verteilung der verschiedenen Topics im Netzwerk scheint keinem besonders augenfälligen Muster zu folgen. Die farblich nach Modularität berechneten Cluster zeigen keine besonderen Häufigkeiten einzelner oder mehrerer Themen. Insgesamt wirkt die Themenverteilung im Netzwerk eher unsystematisch, einzelne Häufungen scheinen jedoch trotzdem erkennbar. So wird der Artikel r12 häufig für „IGP, intraguild predation“ zitiert. Der Artikel r1 wird häufiger als für alle anderen Themen für das Topic „replacement“ angeführt. Insgesamt zeigt sich aber bestätigt, dass Artikel für unterschiedliche Aussagen zu verschiedenen Themen zitiert werden. So stellt es sich auch im Fall des Seed Dokuments

Topics abgrenzen lassen würden. Für die Zielsetzung der vorliegenden Arbeit zeigen sich in dem Netzwerk einige höchst relevante Aspekte invasionsbiologischer Fragestellungen.

4.2.6. Zusammenfassung

Die hier vorgestellten Netzwerke sollten Einsichten in die thematische Multidimensionalität des untersuchten Feldes liefern. Dazu wurden die Zitationen einer Auswahl invasionsbiologischer Artikel mit ihren epistemischen Funktionen kodiert. Anschließend wurden für die Zitationen des Primärnetzwerks Topics ermittelt. Für theoretische Zitationen ließen sich die beiden Felder „competition“ und „distribution“ gegeneinander abgrenzen. Bei beiden handelt es sich um wichtige und klassische Themen der Invasionsbiologie. Methodische Zitationen erscheinen dahingegen recht heterogen. Durch ein höheres Abstraktionsniveau bei der Erhebung der Topics könnten sich ggf. quantitative von qualitativen Untersuchungsmethoden klarer abgrenzen lassen. Recht deutlich im methodischen Netzwerk zu erkennen sind Zitationen, die auf den Ort der Datenerhebung verweisen. Hier deutet sich im vorliegenden Fall ein niederländisches Cluster an. Für die Untersuchung der thematischen Strukturen von auf empirische Objekte bezogenen Zitationen wurden zwei verschiedene Ansätze verfolgt. In dem einen Ansatz wurden die Namen der empirischen Objekte (zumeist Gammaridenarten) als Topics extrahiert, in dem anderen Ansatz wurden deren Eigenschaften aus den Kontexten der Zitationen erhoben. Die Zusammenführung der beiden Ansätze zeigt, dass tendenziell verschiedene Arten für verschiedene Eigenschaften zitiert werden. Im vorliegenden Fall lässt sich daran schon die inhaltliche Ausrichtung des Samples erkennen, da *D. Villosus* (als Invasor, der die heimischen und etablierten Arten bedroht) häufig für sein opportunistisches Fressverhalten zitiert wird und andere Arten (vom Invasor potentiell verdrängt) eher für Attribute wie Anpassungsfähigkeit. Das Netzwerk, das sich aus den auf die Fragestellungen und Forschungsprobleme bezogenen Zitationen erstellen ließ, zeigt einige für die Invasionsbiologie höchst relevante thematische Aspekte. Auch wenn das Netzwerk eine Vielzahl verschiedener Topics präsentiert, so liegen diese doch inhaltlich sehr nah beieinander. Zu erwarten wäre, dass sich die Themenkomplexe im Vergleich mit Netzwerken rein ökologischer oder klassisch biologischer Forschungsarbeiten klar würden abgrenzen lassen.

Zwischen den über die verschiedenen epistemischen Funktionen ermittelten Topics gibt es sowohl Gemeinsamkeiten als auch Unterschiede. So ist das Thema „distribution“, also die Verteilung, bzw. Verbreitung von Arten sowohl in theoretischen Zusammenhängen, als auch innerhalb von Zitationen, die sich auf empirische Objekte oder das Forschungsproblem beziehen relevant. Im Zusammenhang theoretischer Bezüge geht es eher um die grundlegenden Mechanismen der Verbreitung von Arten und um die Faktoren, die diese begünstigen. Wird Verbreitung auf empirische Objekte bezogen, geht es zumeist um das Vorkommen bestimmter Arten in bestimmten Gebieten. Als Forschungsproblem tritt „distribution“ als spezifisches Thema der Invasionsbiologie auf, wenn die Ausbreitung invasiver Arten im Mittelpunkt der Analysen steht. In methodischen Referenzen spielt das Sampling eine Rolle. Bei samplingbezogenen Referenzen geht es zumeist entweder darum den Ort der Datenerhebung unter Zuhilfenahme von Quellen eingehender zu beschreiben oder darum auf die Nachnutzung von Daten zu verweisen, die an einem bestimmten Ort erhoben wurden. Insofern ist die Dimension „distribution“ sogar in methodischen Zitationen zu finden.

Neben „distribution“ ist auch „competition“ ein Thema, dass sich in der ein oder anderen Art in mehreren der erstellten Netzwerke finden lässt. Im theoretischen Netzwerk tritt das Thema alleine auf oder gemeinsam mit „replacement“ oder „impact on benthic communities“. Der Wettbewerb zwischen invasiven und heimischen, bzw. etablierten Arten kann einen starken Einfluss auf das ganze Ökosystem haben und führt in manchen Fällen zur Verdrängung von Arten. Diese Themen finden sich auch im Netzwerk der Zitationen zu Forschungsproblemen. Insbesondere die beiden am häufigsten im Netzwerk auftretenden Topics („replacement“ und „IGP, intraguild predation“) sind inhaltlich mit dem

Begriff „competition“ verknüpft. Konzentriert man sich bei der Erhebung der Topics im Netzwerk für die empirischen Objekte auf die beschriebenen Eigenschaften, zeigt sich das Thema der Konkurrenz zwischen den Arten in Aussagen über ihr Fressverhalten.

Trotz der dargestellten Gemeinsamkeiten der Netzwerke zeigen sie eine thematische Heterogenität hinsichtlich der verschiedenen Zitationsfunktionen. So sind beispielsweise die thematisierten empirischen Objekte allein im entsprechenden Netzwerk vorhanden. Aus den zu den Zitationsbeziehungen ermittelten Topics kann in keinem der anderen Netzwerke das konkrete Untersuchungsobjekt benannt werden. Insofern liefert die Betrachtungsweise einen informativen Mehrwert gegenüber einer Betrachtung die sich etwa ausschließlich auf die in Zitationen thematisierten Forschungsprobleme bezieht. Dies wird vor allem auch im Vergleich zwischen den beiden erprobten Ansätzen (Namen vs. Eigenschaften) zur Erhebung der Topics für empirische Objekte deutlich. Darüber hinaus liefert auch das auf methodischen Zitationen aufbauende Netzwerk einen thematischen Einblick, der so von den anderen Netzwerken nicht ermittelbar ist. Den allgemeinen methodischen Zugang, konkrete verwendete Methoden oder Informationen zur Nachnutzung von Modellparametern oder Daten finden sich ausschließlich im Kontext methodischer Zitationen. Eine getrennte Betrachtung dieser funktionalen Beziehungen scheint daher einen besonderen Mehrwert an Information zu generieren. Der Mehrwert, der sich aus der gesonderten Betrachtung der aus theoretischen Zitationen gewonnenen Topics ergibt, ist dahingegen geringer einzuschätzen. Die Topics, die im theoretischen Zitationsnetzwerk enthalten sind, haben eine große Überschneidung mit den Topics, die sich aus der Betrachtung auf Ebene der Forschungsprobleme ergeben. Insgesamt zeigt die Analyse aber eine thematische Multidimensionalität des Feldes, die mithilfe der getrennten Interpretation nach Zitationsfunktion nachgewiesen werden kann.

5. Diskussion

Für die vorliegende Arbeit wurden eingangs drei Ziele formuliert: (1) die Entwicklung und Dokumentation einer Vorgehensweise zur Berücksichtigung von Zitationsfunktionen bei der Analyse thematischer Strukturen, (2) die vergleichende Analyse der epistemischen Funktionen von Zitationsbeziehungen in einem ausgewählten wissenschaftlichen Feld und (3) die Exploration der Implikationen, die eine Berücksichtigung bzw. Nicht-Berücksichtigung von Zitationsfunktionen auf das bibliometrische Mapping thematischer Strukturen hat. In diesem abschließenden Kapitel soll das entwickelte Vorgehen noch einmal kurz skizziert (bezieht sich auf Ziel Eins) und die empirischen Ergebnisse zusammengefasst werden (bezieht sich auf Ziel Zwei und Drei). Darüber hinaus werden einige Desiderate des vorliegenden Forschungsvorhabens benannt.

Das Forschungsdesign baut grundlegend auf der Entwicklung eines Kategoriensystems zur Klassifizierung von Zitationsfunktionen auf. Das Kategoriensystem wurde in Anlehnung an die in Gläser u. a. (2018) vorgeschlagenen Dimensionen entwickelt und erprobt. Es stellte sich als hilfreich heraus zunächst die zu kodierenden Artikel einmal komplett durchzulesen und jeweils Notizen zu den behandelten Themen der einzelnen Kategorien anzufertigen. Für die Einschätzung der jeweils durch die Zitationen ausgedrückten Funktion wurden sowohl der Kontext, in den die Zitation eingebettet ist, als auch der Inhalt auf den verwiesen wird berücksichtigt. Diese Kombination aus kontext- und inhaltsanalytischem Vorgehen trägt zur validen Erfassung der Zitationsfunktionen bei. Die Verwendung des Datenanalyseprogramms MaxQDA für die Kodierung der Zitationen stellte sich als sehr empfehlenswert heraus. Die übersichtliche Darstellung der Volltexte und kodierten Segmente erleichterte den iterativen Kodierprozess. Darüber hinaus ermöglicht die Software den Export in geeignete Formate zur tiefergehenden Analyse der produzierten Daten.

Die fünf entwickelten Kategorien (research problem, empirical object, method, theoretical, application) stellten sich als hinreichend trennscharf und vollständig heraus. Schwierigkeiten und Unsicherheiten bei der Kodierung traten vor allem zu Beginn des Kodierungsprozesses auf und betrafen hauptsächlich die Unterscheidung zwischen Forschungsproblem und einer der jeweils anderen Kategorien. Die Unsicherheiten bezogen sich häufig auf Segmente in denen das Forschungsproblem direkt einen der anderen Codes betrifft. Bei den anwendungsorientierten Artikeln beispielsweise war die Fragestellung oft sehr auf Management- und Governance-Prozesse fokussiert, sodass die Unterscheidung zwischen „research problem“ und „application“ oft schwer fiel. Ähnliches galt für Artikel vom Typ „Modellierung“ und der Unterscheidung zwischen „research problem“ und „method“. Die intensive Auseinandersetzung mit dem Material und seinen Inhalten führte aber mit der Zeit zu Lerneffekten und Routinisierung der Kodiertätigkeit, sodass später im Prozess die Zuordnung recht bestimmt getroffen werden konnte.

Während die Intrakoderreliabilität durch die Anfertigung von Memos und Notizen, die als Kodierleitfaden dienten und das mehrfache Durchsehen aller vergebenen Codes bestmöglich sichergestellt wurde, wurde sie nicht systematisch getestet. Auch kann über die Interkoderreliabilität an dieser Stelle keine definitive Aussage gemacht werden. Es wurde einmal eine Interpretationssitzung mit Frau Velden (DZHW) durchgeführt, um die Kodierleitfäden zu besprechen und stichprobenartig einige unabhängig

voneinander kodierten Segmente abzugleichen. Zwar ergab diese Diskussion eine hohe Übereinstimmung in der Interpretation, jedoch konnte im Rahmen der Arbeit keine systematische Überprüfung der Intrakoderreliabilität geleistet werden.

Für eine erste Erprobung des Kategoriensystems und um einen generellen Einblick in die Struktur von Zitationsbeziehungen in invasionsbiologischen Zeitschriftenartikeln zu erhalten wurden neun Publikationen für die Analyse ausgewählt. Jeweils drei der Publikationen sind verschiedenen feldinternen Forschungsansätzen zuzuordnen, sodass sie auch hinsichtlich der Unterscheidbarkeit ihrer Zitationsstruktur ausgewertet werden konnten.

Die anschließende Auswertung der Ergebnisse der ersten Phase des empirischen Teils betrifft das zweite definierte Forschungsziel der vorliegenden Arbeit: die vergleichende Analyse der epistemischen Funktionen von Zitationsbeziehungen. Bei der Betrachtung der Häufigkeiten der epistemischen Funktionen ließ sich eine ungleiche Verteilung feststellen. Während das Forschungsproblem betreffende, methodische und auf empirische Objekte bezogene Zitate etwa ähnliche Häufigkeiten aufwiesen, wurden theoretische und anwendungsbezogene Zitationen sehr viel seltener gezählt. Auch unterschieden sich die Verteilungen der Zitationsfunktionen nach feldinternem Forschungsansatz. Wo theoretische und das Forschungsproblem betreffende Zitationen in allen drei Dokumentgruppen gleichermaßen vertreten waren, unterschieden sich Typen invasionsbiologischer Publikationen hinsichtlich der Verwendung anwendungsbezogener, methodischer und auf das empirische Objekt bezogener Quellen. Das Ergebnis unterstreicht die eingangs postulierte Verschiedenheit epistemischer Teilgemeinschaften innerhalb eines Feldes, das durch eine übergreifende Fragestellung, wie im vorliegenden Fall den Invasionsprozessen in der Ökologie, definiert ist. Es konnte gezeigt werden, dass die im Sample repräsentierten Forschungsansätze sich hinsichtlich der Funktion, denen die zitierten Quellen dienen, unterscheiden. Die Ergebnisse legen darüber nahe, dass eine Berücksichtigung der feldinternen Heterogenität von Forschungsansätzen nicht unberücksichtigt bleiben sollte.

Auch verteilen sich die verschiedenen Zitationstypen unterschiedlich auf die einzelnen Kapitel der Publikationen. Es konnte beobachtet werden, dass anwendungsbezogene, das Forschungsproblem betreffende und theoretische Bezüge überwiegend in Einleitung und Diskussion zu finden sind. Außerdem zeigt sich deutlich, dass methodische Referenzen vorrangig in den Methodenkapiteln vorkommen. Bezüge auf das empirische Objekt finden sich mit Ausnahme der Ergebniskapitel in allen Abschnitten. Dieses Muster zeigt sich relativ unabhängig von der Zuordnung der Publikationen zu feldspezifischen Forschungsansätzen. Die Ergebnisse verweisen auf institutionalisierte Zitationspraktiken, die mit hoher Wahrscheinlichkeit so auch in anderen Feldern zu finden sind.

Referenzen, die auf Wissensbestände außerhalb des Feldes der Invasionbiologie verweisen sind zum großen Teil methodischer Natur. Das bedeutet, dass sich die Invasionbiologie, wenn es um die Verwendung von Methoden geht, gerne bei anderen Teildisziplinen bedient. Darüber hinaus sind auch einige Zitationen, die sich auf empirische Objekte oder Forschungsfragen beziehen nicht im Felddatensatz enthalten. Daraus lässt sich folgern, dass auch andere Felder relevante Inhalte zu invasionsbiologischen Fragestellungen und Forschungsobjekten hervorbringen. Theoretische und anwendungsbezogene Referenzen scheinen dahingehend vorrangig innerhalb des Feldes diskutiert zu werden.

Auffallend war, dass sich die zitierten Referenzen nach epistemischer Funktion hinsichtlich ihres Alters unterscheiden. Besonders stark prägt sich dabei der Mittelwertunterschied im Alter der zitierten Literatur von Referenzen mit Bezug auf empirische Objekte gegenüber Referenzen mit Bezug auf das Forschungsproblem aus. Verweise mit Bezug auf das Forschungsproblem sind durchschnittlich um fünf Jahre jünger. Dass in Bezug auf das Forschungsproblem jüngere Publikationen bevorzugt zitiert werden kann entweder auf eine etablierte Praxis oder auf das verhältnismäßig junge Forschungsfeld

zurückgeführt werden. Hier wäre sicher ein Feldvergleich zu etwa geistes- und sozialwissenschaftlichen Fachgemeinschaften vielversprechend und aufschlussreich. Das Vorgehen wie es hier für die Analyse der Zitationsstruktur invasionsbiologischer Publikationen entwickelt wurde, ließe sich prinzipiell aber wohl auf jedes andere Feld übertragen, sodass feldvergleichende Analysen zukünftig realisiert werden könnten.

Neben der deskriptiven Analyse der Zitationsfunktionen (Phase 1) wurde auch deren Einfluss auf die Kartierung thematischer Strukturen wissenschaftlicher Felder untersucht (Phase 2), welches das dritte formulierte Forschungsziel bearbeitet. Dazu wurden ausgehend von einem Seed Dokument Publikationen für die netzwerkanalytische Interpretation ausgewählt. Das mehrstufige Auswahlverfahren des Seed Dokuments folgte dabei einigen theoretischen Überlegungen, beinhaltete aber auch Elemente einer Zufallsauswahl. Das Seed Dokument und alle durch eine direkte Zitationsbeziehung (eingehend oder ausgehend) mit ihm verbundenen Publikationen und die jeweils in den Publikationen enthaltenen Referenzen dienten als Grundlage für die Erstellung der Netzwerke und Herausarbeitung deren thematischer Strukturen. Alle in den Dokumenten enthaltenen und in Web of Science verzeichneten Referenzen wurden mit ihrer Zitationsfunktion nach zuvor entwickeltem und oben erläuterten Schema kodiert. Nach der Kodierung der Zitationsfunktionen wurden Netzwerke erstellt, die jeweils nur die über eine spezifische Funktion verbundenen Publikationen enthalten. Die erstellten Netzwerke geben in dieser Form Auskunft darüber, welche der Publikationen in welcher Funktion mit anderen Publikationen verbunden sind. Es zeigt sich, dass einige Publikationen ausschließlich in einer bestimmten Funktion zitiert werden und andere Publikationen gleich in mehrfacher Art und Weise relevant sind. Für die inhaltliche Analyse sollten die in den Publikationen besprochenen Themen (Topics) herausgearbeitet werden. Dazu wurden die entsprechenden Textstellen innerhalb der Volltexte erneut betrachtet und aus dem Kontext die Inhalte und Themen auf die sich die Zitate beziehen herausgearbeitet. Dabei wurden übergreifende Termini entwickelt, die eine Zusammenführung ähnlicher Aspekte ermöglicht aber inhaltlich Unterschiede weiterhin abbildbar hält. Die herausgearbeiteten Themen wurden in den Netzwerkgrafiken als Kantenbeschriftungen visualisiert. Die Netzwerke wurden dann hinsichtlich ihrer Größe und generellen Struktur, der Anordnung ihrer Knoten und Kanten, der in ihnen besprochenen Themen und der Identifizierbarkeit von Clustern analysiert und interpretiert.

Die einzelnen Netzwerke unterscheiden sich in ihrer Größe und Komplexität. Während die Betrachtung der Funktionen „research problem“ und „empirical object“ recht große und gut vernetzte Netzwerke erzeugten, waren die Netzwerke der Funktionen „theoretical“ und „method“ deutlich kleiner und wiesen weitaus weniger Verbindungen auf. Anwendungsbezogene Zitationen kamen im Sample gar nicht vor, dass eine Analyse der Zitationsfunktion „application“ nicht möglich war. Um die thematischen Strukturen der invasionsbiologischen Publikationen zu untersuchen, wurden die für die einzelnen Netzwerke ermittelten Topics vergleichend analysiert. Auch wenn einige Dimensionen wie beispielsweise die Themen „distribution“ und „competition“ in der einen oder anderen Form in mehreren der erstellten Netzwerke gefunden werden konnten, sind die erstellten Netzwerke überwiegend von thematischer Heterogenität zwischen den verschiedenen Zitationsfunktionen gekennzeichnet. Durch diese Vorgehensweise wurde sichtbar, dass sich die Netzwerke unter Berücksichtigung jeweils nur einer der fünf kategorisierten epistemischen Funktionen sehr deutlich unterscheiden und vor Allem, dass die thematische Struktur innerhalb der einzelnen Netzwerke zum Teil sehr verschieden ist.

Aus dem Netzwerk für die Forschungsobjekte ließen sich die konkreten Namen der jeweils behandelten Arten als Topics extrahieren. Aus dem Netzwerk für methodische Zitationen ließen sich der allgemeinen methodischen Zugang, konkrete verwendete Methoden oder Informationen zur Nachnutzung von Modellparametern und Daten gewinnen. Aus dem Netzwerk für Forschungsprobleme ließen sich einige für die Invasionsbiologie höchst relevante thematische Aspekte ermitteln. Eine getrennte Be-

trachtung dieser drei verschiedenen Netzwerke erbringt daher einen besonderen inhaltlichen Mehrwert. Die Topics, die im theoretischen Zitationsnetzwerk enthalten sind, haben eine große Überschneidung mit den Topics, die sich aus der Betrachtung auf Ebene der Forschungsprobleme ergeben, weshalb für sie für die Erhebung relevanter Themen als weniger relevant eingestuft wird. Die Unterscheidung zwischen Forschungsproblem, epistemischem Objekt und Methode scheint für die Erhebung aller relevanten Themen und deren strukturellen Verbindungen auszureichen. Insgesamt zeigt die Analyse aber eine thematische Multidimensionalität des Feldes, die mithilfe der getrennten Interpretation nach Zitationsfunktion nachgewiesen werden kann.

Da sich die Zusammensetzung des Netzwerkes auf die Auswahl eines einzigen Seed Dokuments begründet, ergab sich ein inhaltlich sehr homogenes Sample. Interessant wäre daher sicher die Auswertung eines Netzwerkes das eine größere inhaltliche Heterogenität abbildet. Die Abgrenzung von themenspezifischen Clustern könnte sich in einem komplexen und thematisch heterogenen Netzwerk ggf. deutlicher darstellen. Die Auswahl der Publikationen könnte dafür etwa auf einer lexikalischen Abfrage basieren oder auf der Auswahl einer Fachzeitschrift. Statt ausgehend von einem einzigen Seed Dokument das Netzwerk aufzubauen, könnten auch mehrere Publikationen und deren ein- und ausgehende Referenzen die Grundlage der Netzwerkanalyse bilden. Bei komplexeren Netzwerken könnte sich auch die Anwendung automatisierter Modelle zur Topic Extraction als interessante Alternative zur manuellen, das heißt intellektuellen Erschließung der Topics herausstellen. Da die intellektuelle Erschließung der Topics aus den Kontexten der Zitationen wie sie in der vorliegenden Arbeit durchgeführt wurde auch keine systematische Überprüfung auf Inter- und Intrakoderreliabilität erfahren hat, wäre es darüber hinaus interessant das Ergebnis des Verfahrens mit den Ergebnissen automatisierter Verfahren zu vergleichen.

In der vorliegenden Arbeit sollte nicht zuletzt überhaupt eine Vorgehensweise zur Berücksichtigung von Zitationsfunktionen bei der Analyse thematischer Strukturen von Zitationsnetzwerken entwickelt werden (erstes Ziel). Das an dieser Stelle formulierte Resümee zeigt zum einen wie eine solche Berücksichtigung ermöglicht werden kann, welche Desiderate das Forschungsdesign hinterlässt und welche Ergebnisse aus der Analyse gewonnen werden konnten. Abschließend lässt sich festhalten, dass das in der vorliegenden Arbeit entwickelte Forschungsdesign für die Beantwortung des eingangs formulierten Forschungsfragen geeignet ist (zweites Ziel) und einige interessante Einsichten zu den epistemischen Funktionen von Zitierungen und deren Implikationen für die Modellierung von Zitationsnetzwerken beiträgt (drittes Ziel). Damit schließt die Forschungsarbeit zum einen an aktuelle Diskurse zur Identifizierung und Abgrenzung disziplinärer Felder und der in ihnen behandelten thematischen Strukturen als auch an die theoretische und empirische Klassifikation von Zitationsfunktionen und deren Potential zur Beschreibung wissenschaftlicher Felder an. Im größeren Kontext trägt sie somit zum besseren Verständnis wissenschaftlicher Praxis und insbesondere den Mechanismen der Produktion wissenschaftlichen Wissens bei.

6. Literaturverzeichnis

- Bastian, Mathieu, Sebastien Heymann und Mathieu Jacomy (2009). *Gephi: an open source software for exploring and manipulating networks: International AAAI Conference on Weblogs and Social Media*.
- Blei, David M. (2012a). „Probabilistic topic models“. In: *Communications of the ACM* 55.4, S. 77–84.
- (2012b). „Probabilistic topic models“. In: *Communications of the ACM* 55.4, S. 77.
- Blei, David M., Andrew Y. Ng und Michael I. Jordan (2003). „Latent Dirichlet Allocation“. In: *Journal of Machine Learning Research* 3, S. 993–1022.
- Bornmann, Lutz und Hans-Dieter Daniel (2008). „What do citation counts measure? A review of studies on citing behavior“. In: *Journal of documentation*.
- Boyack, Kevin W., Richard Klavans und Katy Börner (2005). „Mapping the backbone of science“. In: *Scientometrics* 64.3, S. 351–374.
- Brooks, T. A. (1986). „Evidence of complex citer motivations“. In: *Journal of the American Society for Information Science* 37, S. 34–36.
- Brooks, Terrence A. (1985). „Private acts and public objects – an investigation of citer motivations“. In: *Journal of the American Society for Information Science* 36, S. 223–229.
- Callon, Michel, Jean-Pierre Courtial, William A. Turner und Serge Bauin (1983). „From translations to problematic networks: An introduction to co-word analysis“. In: *Information (International Social Science Council)* 22.2, S. 191–235.
- Cano, V. (1989). „Citation behavior: classification, utility, and location“. In: *Journal of the American Society for Information Science* 40, S. 284–290.
- Chubin, Daryl E. und Soumyo D. Moitra (1975). „Content Analysis of References: Adjunct or Alternative to Citation Counting?“. In: *Social studies of science* 5.4, S. 423–441.
- Cole, Stephen (1975). „The growth of scientific knowledge: theories of deviance as a case study“. In: *The idea of social structure: Papers in honor of Robert K. Merton*. Hrsg. von Lewis A. Coser. New York: Harcourt Brace Jovanovich.
- Cronin, Blaise (1984). *The Citation Process: The role and significance of citations in scientific communication*. Oxford: Taylor Graham.
- Davis, Mark A. (2006). „Invasion biology 1958-2005: the pursuit of science and conservation“. In: *Conceptual Ecology and Invasion Biology: Reciprocal Approaches to Nature*. Hrsg. von Marc William Cadotte, Sean M. McMahon und Tadashi Fukami. Dordrecht: Kluwer Academic Publishers, S. 35–64.
- Fleck, Ludwik (2012 [1935]). *Entstehung und Entwicklung einer wissenschaftlichen Tatsache: Einführung in die Lehre vom Denkstil und Denkkollektiv*. Mit einer Einleitung herausgegeben von Lothar Schäfer und Thomas Schnelle, 9. Auflage. Frankfurt am Main: Suhrkamp.

- Garfield, Eugene (1962). „Can Citation Indexing be Automated?“ In: *Essays of an Information Scientist* Vol. 1, S. 84–90.
- (1978). „High impact science and the case of Arthur Jensen“. In: *Current Contents* 41, S. 5–15.
- Garfield, Eugene und Alfred Welljams-Dorof (1990). „The impact of fraudulent research on the scientific literature: The Stephen E. Breuning case“. In: *Jama* 263.10, S. 1424–1426.
- Glänzel, Wolfgang und Hans-Jürgen Czerwon (1995). „A New Methodological Approach to Bibliographic Coupling and Its Application to Research-Front and Other Core Documents“. In: *Proceedings of 5th International Conference on Scientometrics and Informetrics*, S. 167–176.
- (1996). „A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level“. In: *Scientometrics* 37.2, S. 195–221.
- Gläser, Jochen (2006). *Wissenschaftliche Produktionsgemeinschaften: Die soziale Ordnung der Forschung*. Frankfurt/New York: Campus Verlag.
- Gläser, Jochen, Wolfgang Glänzel und Andrea Scharnhorst (2017). „Same data—different results? Towards a comparative approach to the identification of thematic structures in science“. In: *Scientometrics* 111.2, S. 981–998.
- Gläser, Jochen, Grit Laudel, Christopher Grieser und Uli Meyer (2018). „Scientific fields as epistemic regimes: new opportunities for comparative science studies“. In: *Technical University Technology Studies Working Papers TUTS-WP-3-2018*.
- Glenisson, Patrick, Wolfgang Glänzel, Frizo Janssens und Bart de Moor (2005). „Combining full text and bibliometric information in mapping scientific disciplines“. In: *Information processing & management* 41.6, S. 1548–1572.
- Glenisson, Patrick, Wolfgang Glänzel und Olle Persson (2005). „Combining full-text analysis and bibliometric indicators. A pilot study“. In: *Katholieke Universiteit Leuven, Open Access publications from Katholieke Universiteit Leuven* 63.
- Griffith, Belver C., Henry G. Small, Judith A. Stonehill und Sandra Dey (1974). „The structure of scientific literatures II: Toward a macro-and microstructure for science“. In: *Science studies* 4.4, S. 339–365.
- Hooten, Patricia A. (1991). „Frequency and functional use of cited documents in information science“. In: *Journal of the American Society for Information Science* 42.6, S. 397–404.
- Jacomy, Mathieu, Tommaso Venturini, Sebastien Heymann und Mathieu Bastian (2014). „ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software“. In: *PloS one* 9.6, e98679.
- Janssens, Frizo, Wolfgang Glänzel und Bart de Moor (2007). „Dynamic hybrid clustering of bioinformatics by incorporating text mining and citation analysis“. In: *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*. San Jose, California, USA, S. 360–369.
- Jarneving, Bo (2001). „The cognitive structure of current cardiovascular research“. In: *Scientometrics* 50.3, S. 365–389.
- Kessler, Maxwell Mirton (1963). „Bibliographic coupling between scientific papers“. In: *American Documentation* 14.1, S. 10–25.

- Kirschner, Heiko (2014). „Karin Knorr Cetina: Von der Fabrikation von Erkenntnis zu Wissenskulturen“. In: *Schlüsselwerke der Science & Technology Studies*. Hrsg. von Diana Lengersdorf und Matthias Wieser. Wiesbaden: Springer VS, S. 123–132.
- Klavans, Richard und Kevin W. Boyack (2011). „Using global mapping to create more accurate document-level maps of research fields“. In: *Journal of the American Society for Information Science and Technology* 62.1, S. 1–18.
- Knorr Cetina, Karin (1981). *The manufacture of knowledge : an essay on the constructivist and contextual nature of science*.
- (1999). *Epistemic Cultures: How the Sciences make Knowledge*. Cambridge (Mass.)/London: Harvard University Press.
 - (2002 [1999]). *Wissenskulturen: Ein Vergleich naturwissenschaftlicher Wissensformen*. 1. Auflage. Frankfurt am Main: Suhrkamp.
- Kochan, Carol Ann und John M. Budd (1992). „The persistence of fraud in the literature: the Darsee case“. In: *Journal of the American Society for Information Science* 43.7, S. 488–493.
- Krampen, Günter, Leo Montada und Patrik Burkard (2002). *Wissenschaftsforschung in der Psychologie*. Göttingen: Hogrefe Verl. für Psychologie.
- Kuhn, Thomas S. (2014 [1962]). *Die Struktur wissenschaftlicher Revolutionen*. Zweite revidierte und um das Postskriptum von 1969 ergänzte Auflage, 24. Auflage. Suhrkamp-Taschenbuch Wissenschaft. Frankfurt am Main: Suhrkamp.
- Latour, Bruno und Steve Woolgar (1986 [1979]). *Laboratory Life: The Construction of Scientific Facts*. Introduction by Jonas Salk, with a new postscript and index by the authors. Princeton, New Jersey: Princeton University Press.
- Lengersdorf, Diana und Matthias Wieser, Hrsg. (2014). *Schlüsselwerke der Science & Technology Studies*. Wiesbaden: Springer VS.
- Leydesdorff, Loet (1986). „The development of frames of references“. In: *Scientometrics* 9.3-4, S. 103–125.
- (2004). „Clusters and maps of science journals based on bi-connected graphs in journal citation reports“. In: *Journal of documentation*.
- Leydesdorff, Loet und Susan Cozzens (1993). „The delineation of specialties in terms of journals using the dynamic journal set of the SCI“. In: *Scientometrics* 26.1, S. 135–156.
- Leydesdorff, Loet und Ismael Rafols (2009). „A global map of science based on the ISI subject categories“. In: *Journal of the American Society for Information Science and Technology* 60.2, S. 348–362.
- Leydesdorff, Loet, Paul Wouters und Lutz Bornmann (2016). „Professional and citizen bibliometrics: complementarities and ambivalences in the development and use of indicators—a state-of-the-art report“. In: *Scientometrics* 109.3, S. 2129–2150.
- Lipetz, Ben-Ami (1965). „Improvement of the selectivity of citation indexes to science literature through inclusion of citation relationship indicators“. In: *American Documentation* 16.2, S. 81–90.
- Maričić, Siniša, Jagoda Spaventi, Leo Pavičić und Greta Pifat-Mrzljak (1998). „Citation context versus the frequency counts of citation histories“. In: *Journal of the American Society for Information Science* 49.6, S. 530–540.

- Mayntz, Renate und Fritz W. Scharpf (1995). „Der Ansatz des akteurzentrierten Institutionalismus“. In: *Gesellschaftliche Selbstregulierung und politische Steuerung*. Hrsg. von Renate Mayntz und Fritz W. Scharpf. Frankfurt am Main: Campus, S. 39–72.
- McCain, Katherine W. und Kathleen Turner (1989). „Citation context analysis and aging patterns of journal articles in molecular genetics“. In: *Scientometrics* 17.1-2, S. 127–163.
- Merton, Robert K. (1942). „A Note on Science and Democracy“. In: *Journal of Legal and Political Sociology* 1, S. 115–126.
- Moravcsik, Michael J. und Poovanalingam Murugesan (1975). „Some results on the function and quality of citations“. In: *Social studies of science* 5.1, S. 86–92.
- Murugesan, Poovanalingam und Michael J. Moravcsik (1978). „Variation of the nature of citation measures with journals and scientific specialties“. In: *Journal of the American Society for Information Science* 29.3, S. 141–147.
- Oppenheim, Charles und Susan P. Renn (1978). „Highly cited old papers and the reasons why they continue to be cited“. In: *Journal of the American Society for Information Science* 29.5, S. 225–231.
- Panofsky, Aaron L. (2010). „A Critical Reconsideration of the Ethos and Autonomy of Science“. In: *Robert K. Merton*. Hrsg. von Craig J. Calhoun. A Columbia / SSRC book. New York: Columbia University Press, S. 140–163.
- Peritz, Bluma C. (1983). „A classification of citation roles for the social sciences and related fields“. In: *Scientometrics* 5.5, S. 303–312.
- R Core Team (2017). „R. A language and environment for statistical computing: R Foundation for Statistical Computing, Vienna, Austria“. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Rip, Arie und Jean-Pierre Courtial (1984). „Co-word maps of biotechnology: An example of cognitive scientometrics“. In: *Scientometrics* 6.6, S. 381–400.
- Schiminovich, Samuel (1971). „Automatic classification and retrieval of documents by means of a bibliographic pattern discovery algorithm“. In: *Information Storage and Retrieval* 6.6, S. 417–435.
- Shadish, William R., Donna Tolliver, Maria Gray und Sunil K. Sen Gupta (1995). „Author judgments about works they cite: three studies from psychology journals“. In: *Social studies of science* 25, S. 477–498.
- Small, Henry G. (1973). „Co-citation in the scientific literature: A new measure of the relationship between two documents“. In: *Journal of the American Society for Information Science* 24.4, S. 265–269.
- (1977). „A co-citation model of a scientific specialty: A longitudinal study of collagen research“. In: *Social studies of science* 7.2, S. 139–166.
- Small, Henry und Belver C. Griffith (1974). „The structure of scientific literatures I: Identifying and graphing specialties“. In: *Science studies* 4.1, S. 17–40.
- Snyder, Herbert und Susan Bonzi (1998). „Patterns of self-citation across disciplines (1980–1989)“. In: *Journal of Information Science* 24, S. 431–435.
- Spiegel-Rösing, Ina (1977). „Science Studies: Bibliometric and Content Analysis“. In: *Social studies of science* 7.1, S. 97–113.

- Tahamtan, Iman und Lutz Bornmann (2018). „Core elements in the process of citing publications: Conceptual overview of the literature“. In: *Journal of Informetrics* 12.1, S. 203–216.
- (2019). „What do citation counts measure? An updated review of studies on citations in scientific documents published between 2006 and 2018“. In: *Scientometrics* 121.3, S. 1635–1684.
- Tijssen, Robert J. W. (1992). „A quantitative assessment of interdisciplinary structures in science and technology: co-classification analysis of energy research“. In: *Research Policy* 21.1, S. 27–44.
- VERBI Software. *MAXQDA 2020: Software für qualitative Datenanalyse*. Berlin, Deutschland.
- Vinkler, Peter (1987). „A quasi-quantitative citation model“. In: *Scientometrics* 12, S. 47–72.
- Waltman, Ludo und Nees Jan van Eck (2012). „A new methodology for constructing a publication-level classification system of science“. In: *Journal of the American Society for Information Science and Technology* 63.12, S. 2378–2392.
- White, Howard D. und Belver C. Griffith (1981). „Author cocitation: A literature measure of intellectual structure“. In: *Journal of the American Society for Information Science* 32.3, S. 163–171.
- White, Howard D. und Katherine W. McCain (1998). „Visualizing a discipline: An author co-citation analysis of information science, 1972–1995“. In: *Journal of the American Society for Information Science* 49.4, S. 327–355.
- Yau, Chyi-Kwei, Alan Porter, Nils Newman und Arho Suominen (2014). „Clustering scientific documents with topic modeling“. In: *Scientometrics* 100.3, S. 767–786.
- van Raan, Anthony und Robert Tijssen (1993). „The neural net of neural network research: An exercise in bibliometric mapping“. In: *Scientometrics* 26.1, S. 169–192.

A. Anhang – Dokumentation und Material zum Sampling

A.1. Dokumentation Lexical Query Felddatensatz Invasionsbiologie

#1 OR #2 OR #3 OR #4 OR #5 OR #6

Query Part #1

((TS=(„invasive population“ OR „invasive range“ OR „invasive invader“ OR „invasive spread“ OR „invasive community“ OR „invasive genotype“ OR „invasive prey“ OR „invasive habitat“ OR „invasive abundance“ OR „alien population“ OR „alien range“ OR „alien invader“ OR „alien spread“ OR „alien community“ OR „alien genotype“ OR „alien prey“ OR „alien habitat“ OR „alien abundance“ OR „introduced population“ OR „introduced range“ OR „introduced invader“ OR „introduced spread“ OR „introduced community“ OR „introduced genotype“ OR „introduced prey“ OR „introduced habitat“ OR „introduced abundance“ OR „non-native population“ OR „non-native range“ OR „non-native invader“ OR „non-native spread“ OR „non-native community“ OR „non-native genotype“ OR „non-native prey“ OR „non-native habitat“ OR „non-native abundance“ OR „nonnative population“ OR „nonnative range“ OR „nonnative invader“ OR „nonnative spread“ OR „nonnative community“ OR „nonnative genotype“ OR „nonnative prey“ OR „nonnative habitat“ OR „nonnative abundance“ OR „nonindigenous population“ OR „nonindigenous range“ OR „nonindigenous invader“ OR „nonindigenous spread“ OR „nonindigenous community“ OR „nonindigenous genotype“ OR „nonindigenous prey“ OR „nonindigenous habitat“ OR „nonindigenous abundance“ OR „non-indigenous population“ OR „non-indigenous range“ OR „non-indigenous invader“ OR „non-indigenous spread“ OR „non-indigenous community“ OR „non-indigenous genotype“ OR „non-indigenous prey“ OR „non-indigenous habitat“ OR „non-indigenous abundance“ OR „allochthonous population“ OR „allochthonous range“ OR „allochthonous invader“ OR „allochthonous spread“ OR „allochthonous community“ OR „allochthonous genotype“ OR „allochthonous prey“ OR „allochthonous habitat“ OR „allochthonous abundance“ OR „exotic population“ OR „exotic range“ OR „exotic invader“ OR „exotic spread“ OR „exotic community“ OR „exotic genotype“ OR „exotic prey“ OR „exotic habitat“ OR „exotic abundance“ OR „invasion history“ OR „invasion histories“ OR „invasion dynamic“ OR „invasion impact“ OR „invasion management“ OR „invasion pathway“ OR „invasion resistance“ OR „habitat invasion ×
))) AND DOCUMENT TYPES: (Article OR Review)
Indexes=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH

Query Part #2

((TS=(„Ecological invasion“ or „Biological invasion“ or „Invasion biology“ or „Invasion ecology“ or (invasive NEAR/2 species) or (alien NEAR/2 species) or („introduced“ NEAR/2 species) or (non-

native NEAR/2 species) or (nonnative NEAR/2 species) or (nonindigenous NEAR/2 species) or (non-indigenous NEAR/2 species) or (allochthonous NEAR/2 species) or (exotic NEAR/2 species)))) AND DOCUMENT TYPES: (Article OR Review)
Indexes=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH

Query Part #3

((TS=((invasive NEAR/2 plant) OR (invasive NEAR/2 fish) OR (invasive NEAR/2 pest) OR (invasive NEAR/2 tree) OR (invasive NEAR/2 weed) OR (invasive NEAR/2 predator) OR (invasive NEAR/2 ant) OR (invasive NEAR/2 insect) OR (invasive NEAR/2 shrub) OR (invasive NEAR/2 crayfish) OR (invasive NEAR/2 mussel) OR (invasive NEAR/2 taxon) OR (invasive NEAR/2 pathogen) OR (invasive NEAR/2 plantation) OR (invasive NEAR/2 mammal) OR (invasive NEAR/2 earthworm) OR (invasive NEAR/2 forest) OR (invasive NEAR/2 algae) OR (invasive NEAR/2 trout) OR (invasive NEAR/2 herbivore) OR (invasive NEAR/2 snail) OR (invasive NEAR/2 beetle) OR (invasive NEAR/2 bivalve) OR (invasive NEAR/2 forb) OR (invasive NEAR/2 macrophyte) OR (invasive NEAR/2 bird) OR (invasive NEAR/2 crab) OR (invasive NEAR/2 pine) OR (invasive NEAR/2 vegetation) OR (invasive NEAR/2 animal) OR (invasive NEAR/2 goby) OR (invasive NEAR/2 mosquito) OR (invasive NEAR/2 parasitoid) OR (invasive NEAR/2 rat) OR (invasive NEAR/2 acacia) OR (invasive NEAR/2 lionfish) OR (invasive NEAR/2 amphipod) OR (invasive NEAR/2 seaweed) OR (invasive NEAR/2 grassland) OR (invasive NEAR/2 parasite) OR (invasive NEAR/2 carp) OR (invasive NEAR/2 seed) OR (invasive NEAR/2 herb) OR (invasive NEAR/2 invertebrate) OR (invasive NEAR/2 „bee“) OR (invasive NEAR/2 rodent) OR (invasive NEAR/2 vertebrate) OR (invasive NEAR/2 salmonid) OR (invasive NEAR/2 „vine“) OR (invasive NEAR/2 congener) OR (invasive NEAR/2 ascidian) OR (invasive NEAR/2 flora) OR (invasive NEAR/2 toad) OR (invasive NEAR/2 spartina) OR (alien NEAR/2 plant) OR (alien NEAR/2 fish) OR (alien NEAR/2 pest) OR (alien NEAR/2 tree) OR (alien NEAR/2 weed) OR (alien NEAR/2 predator) OR (alien NEAR/2 ant) OR (alien NEAR/2 insect) OR (alien NEAR/2 shrub) OR (alien NEAR/2 crayfish) OR (alien NEAR/2 mussel) OR (alien NEAR/2 taxon) OR (alien NEAR/2 pathogen) OR (alien NEAR/2 plantation) OR (alien NEAR/2 mammal) OR (alien NEAR/2 earthworm) OR (alien NEAR/2 forest) OR (alien NEAR/2 algae) OR (alien NEAR/2 trout) OR (alien NEAR/2 herbivore) OR (alien NEAR/2 snail) OR (alien NEAR/2 beetle) OR (alien NEAR/2 bivalve) OR (alien NEAR/2 forb) OR (alien NEAR/2 macrophyte) OR (alien NEAR/2 bird) OR (alien NEAR/2 crab) OR (alien NEAR/2 pine) OR (alien NEAR/2 vegetation) OR (alien NEAR/2 animal) OR (alien NEAR/2 goby) OR (alien NEAR/2 mosquito) OR (alien NEAR/2 parasitoid) OR (alien NEAR/2 rat) OR (alien NEAR/2 acacia) OR (alien NEAR/2 lionfish) OR (alien NEAR/2 amphipod) OR (alien NEAR/2 seaweed) OR (alien NEAR/2 grassland) OR (alien NEAR/2 parasite) OR (alien NEAR/2 carp) OR (alien NEAR/2 seed) OR (alien NEAR/2 herb) OR (alien NEAR/2 invertebrate) OR (alien NEAR/2 „bee“) OR (alien NEAR/2 rodent) OR (alien NEAR/2 vertebrate) OR (alien NEAR/2 salmonid) OR (alien NEAR/2 „vine“) OR (alien NEAR/2 congener) OR (alien NEAR/2 ascidian) OR (alien NEAR/2 flora) OR (alien NEAR/2 toad) OR (alien NEAR/2 spartina)))) AND DOCUMENT TYPES: (Article OR Review)
Indexes=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH

Query Part #4

((TS=((„introduced“ NEAR/2 plant) OR („introduced“ NEAR/2 fish) OR („introduced“ NEAR/2 pest) OR („introduced“ NEAR/2 tree) OR („introduced“ NEAR/2 weed) OR („introduced“ NEAR/2 predator) OR („introduced“ NEAR/2 ant) OR („introduced“ NEAR/2 insect) OR („intro-

duced“ NEAR/2 shrub) OR („introduced“ NEAR/2 crayfish) OR („introduced“ NEAR/2 mussel) OR („introduced“ NEAR/2 taxon) OR („introduced“ NEAR/2 pathogen) OR („introduced“ NEAR/2 plantation) OR („introduced“ NEAR/2 mammal) OR („introduced“ NEAR/2 earthworm) OR („introduced“ NEAR/2 forest) OR („introduced“ NEAR/2 algae) OR („introduced“ NEAR/2 trout) OR („introduced“ NEAR/2 herbivore) OR („introduced“ NEAR/2 snail) OR („introduced“ NEAR/2 beetle) OR („introduced“ NEAR/2 bivalve) OR („introduced“ NEAR/2 forb) OR („introduced“ NEAR/2 macrophyte) OR („introduced“ NEAR/2 bird) OR („introduced“ NEAR/2 crab) OR („introduced“ NEAR/2 pine) OR („introduced“ NEAR/2 vegetation) OR („introduced“ NEAR/2 animal) OR („introduced“ NEAR/2 goby) OR („introduced“ NEAR/2 mosquito) OR („introduced“ NEAR/2 parasitoid) OR („introduced“ NEAR/2 rat) OR („introduced“ NEAR/2 acacia) OR („introduced“ NEAR/2 lionfish) OR („introduced“ NEAR/2 amphipod) OR („introduced“ NEAR/2 seaweed) OR („introduced“ NEAR/2 grassland) OR („introduced“ NEAR/2 parasite) OR („introduced“ NEAR/2 carp) OR („introduced“ NEAR/2 seed) OR („introduced“ NEAR/2 herb) OR („introduced“ NEAR/2 invertebrate) OR („introduced“ NEAR/2 „bee“) OR („introduced“ NEAR/2 rodent) OR („introduced“ NEAR/2 vertebrate) OR („introduced“ NEAR/2 salmonid) OR („introduced“ NEAR/2 „vine“) OR („introduced“ NEAR/2 congener) OR („introduced“ NEAR/2 ascidian) OR („introduced“ NEAR/2 flora) OR („introduced“ NEAR/2 toad) OR („introduced“ NEAR/2 spartina) OR (non-native NEAR/2 plant) OR (non-native NEAR/2 fish) OR (non-native NEAR/2 pest) OR (non-native NEAR/2 tree) OR (non-native NEAR/2 weed) OR (non-native NEAR/2 predator) OR (non-native NEAR/2 ant) OR (non-native NEAR/2 insect) OR (non-native NEAR/2 shrub) OR (non-native NEAR/2 crayfish) OR (non-native NEAR/2 mussel) OR (non-native NEAR/2 taxon) OR (non-native NEAR/2 pathogen) OR (non-native NEAR/2 plantation) OR (non-native NEAR/2 mammal) OR (non-native NEAR/2 earthworm) OR (non-native NEAR/2 forest) OR (non-native NEAR/2 algae) OR (non-native NEAR/2 trout) OR (non-native NEAR/2 herbivore) OR (non-native NEAR/2 snail) OR (non-native NEAR/2 beetle) OR (non-native NEAR/2 bivalve) OR (non-native NEAR/2 forb ×) OR (non-native NEAR/2 macrophyte) OR (non-native NEAR/2 bird) OR (non-native NEAR/2 crab) OR (non-native NEAR/2 pine) OR (non-native NEAR/2 vegetation) OR (non-native NEAR/2 animal) OR (non-native NEAR/2 goby) OR (non-native NEAR/2 mosquito) OR (non-native NEAR/2 parasitoid) OR (non-native NEAR/2 rat) OR (non-native NEAR/2 acacia) OR (non-native NEAR/2 lionfish) OR (non-native NEAR/2 amphipod) OR (non-native NEAR/2 seaweed) OR (non-native NEAR/2 grassland) OR (non-native NEAR/2 parasite) OR (non-native NEAR/2 carp) OR (non-native NEAR/2 seed) OR (non-native NEAR/2 herb) OR (non-native NEAR/2 invertebrate) OR (non-native NEAR/2 „bee“) OR (non-native NEAR/2 rodent) OR (non-native NEAR/2 vertebrate) OR (non-native NEAR/2 salmonid) OR (non-native NEAR/2 „vine“) OR (non-native NEAR/2 congener) OR (non-native NEAR/2 ascidian) OR (non-native NEAR/2 flora) OR (non-native NEAR/2 toad) OR (non-native NEAR/2 spartina)))) AND DOCUMENT TYPES: (Article OR Review)

Indexes=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH

Query Part #5

((TS=((nonnative NEAR/2 plant) OR (nonnative NEAR/2 fish) OR (nonnative NEAR/2 pest) OR (nonnative NEAR/2 tree) OR (nonnative NEAR/2 weed) OR (nonnative NEAR/2 predator) OR (nonnative NEAR/2 ant) OR (nonnative NEAR/2 insect) OR (nonnative NEAR/2 shrub) OR (nonnative NEAR/2 crayfish) OR (nonnative NEAR/2 mussel) OR (nonnative NEAR/2 taxon) OR (nonnative NEAR/2 pathogen) OR (nonnative NEAR/2 plantation) OR (nonnative NEAR/2 mammal) OR (nonnative NEAR/2 earthworm) OR (nonnative NEAR/2 forest) OR (nonnative NEAR/2 algae) OR (nonnative NEAR/2 trout) OR (nonnative NEAR/2 herbivore) OR (nonnative NEAR/2

snail) OR (nonnative NEAR/2 beetle) OR (nonnative NEAR/2 bivalve) OR (nonnative NEAR/2 forb) OR (nonnative NEAR/2 macrophyte) OR (nonnative NEAR/2 bird) OR (nonnative NEAR/2 crab) OR (nonnative NEAR/2 pine) OR (nonnative NEAR/2 vegetation) OR (nonnative NEAR/2 animal) OR (nonnative NEAR/2 goby) OR (nonnative NEAR/2 mosquito) OR (nonnative NEAR/2 parasitoid) OR (nonnative NEAR/2 rat) OR (nonnative NEAR/2 acacia) OR (nonnative NEAR/2 lionfish) OR (nonnative NEAR/2 amphipod) OR (nonnative NEAR/2 seaweed) OR (nonnative NEAR/2 grassland) OR (nonnative NEAR/2 parasite) OR (nonnative NEAR/2 carp) OR (nonnative NEAR/2 seed) OR (nonnative NEAR/2 herb) OR (nonnative NEAR/2 invertebrate) OR (nonnative NEAR/2 „bee“) OR (nonnative NEAR/2 rodent) OR (nonnative NEAR/2 vertebrate) OR (nonnative NEAR/2 salmonid) OR (nonnative NEAR/2 „vine“) OR (nonnative NEAR/2 congener) OR (nonnative NEAR/2 ascidian) OR (nonnative NEAR/2 flora) OR (nonnative NEAR/2 toad) OR (nonnative NEAR/2 spartina) OR (nonindigenous NEAR/2 plant) OR (nonindigenous NEAR/2 fish) OR (nonindigenous NEAR/2 pest) OR (nonindigenous NEAR/2 tree) OR (nonindigenous NEAR/2 weed) OR (nonindigenous NEAR/2 predator) OR (nonindigenous NEAR/2 ant) OR (nonindigenous NEAR/2 insect) OR (nonindigenous NEAR/2 shrub) OR (nonindigenous NEAR/2 crayfish) OR (nonindigenous NEAR/2 mussel) OR (nonindigenous NEAR/2 taxon) OR (nonindigenous NEAR/2 pathogen) OR (nonindigenous NEAR/2 plantation) OR (nonindigenous NEAR/2 mammal) OR (nonindigenous NEAR/2 earthworm) OR (nonindigenous NEAR/2 forest) OR (nonindigenous NEAR/2 algae) OR (nonindigenous NEAR/2 trout) OR (nonindigenous NEAR/2 herbivore) OR (nonindigenous NEAR/2 snail) OR (nonindigenous NEAR/2 beetle) OR (nonindigenous NEAR/2 bivalve) OR (nonindigenous NEAR/2 forb) OR (nonindigenous NEAR/2 macrophyte) OR (nonindigenous NEAR/2 bird) OR (nonindigenous NEAR/2 crab) OR (nonindigenous NEAR/2 pine) OR (nonindigenous NEAR/2 vegetation) OR (nonindigenous NEAR/2 animal) OR (nonindigenous NEAR/2 goby) OR (nonindigenous NEAR/2 mosquito) OR (nonindigenous NEAR/2 parasitoid) OR (nonindigenous NEAR/2 rat) OR (nonindigenous NEAR/2 acacia) OR (nonindigenous NEAR/2 lionfish) OR (nonindigenous NEAR/2 amphipod) OR (nonindigenous NEAR/2 seaweed) OR (nonindigenous NEAR/2 grassland) OR (nonindigenous NEAR/2 parasite) OR (nonindigenous NEAR/2 carp) OR (nonindigenous NEAR/2 seed) OR (nonindigenous NEAR/2 herb) OR (nonindigenous NEAR/2 invertebrate) OR (nonindigenous NEAR/2 „bee“) OR (nonindigenous NEAR/2 rodent) OR (nonindigenous NEAR/2 vertebrate) OR (nonindigenous NEAR/2 salmonid) OR (nonindigenous NEAR/2 „vine“) OR (nonindigenous NEAR/2 congener) OR (nonindigenous NEAR/2 ascidian) OR (nonindigenous NEAR/2 flora) OR (nonindigenous NEAR/2 toad) OR (nonindigenous NEAR/2 spartina)))) AND DOCUMENT TYPES: (Article OR Review)
Indexes=SCI-EXPANDED, SSCI, CPCI-S, CPCI-SSH

Query Part #6

((TS=((non-indigenous NEAR/2 plant) OR (non-indigenous NEAR/2 fish) OR (non-indigenous NEAR/2 pest) OR (non-indigenous NEAR/2 tree) OR (non-indigenous NEAR/2 weed) OR (non-indigenous NEAR/2 predator) OR (non-indigenous NEAR/2 ant) OR (non-indigenous NEAR/2 insect) OR (non-indigenous NEAR/2 shrub) OR (non-indigenous NEAR/2 crayfish) OR (non-indigenous NEAR/2 mussel) OR (non-indigenous NEAR/2 taxon) OR (non-indigenous NEAR/2 pathogen) OR (non-indigenous NEAR/2 plantation) OR (non-indigenous NEAR/2 mammal) OR (non-indigenous NEAR/2 earthworm) OR (non-indigenous NEAR/2 forest) OR (non-indigenous NEAR/2 algae) OR (non-indigenous NEAR/2 trout) OR (non-indigenous NEAR/2 herbivore) OR (non-indigenous NEAR/2 snail) OR (non-indigenous NEAR/2 beetle) OR (non-indigenous NEAR/2 bivalve) OR (non-indigenous NEAR/2 forb) OR (non-indigenous NEAR/2 macrophyte) OR (non-indigenous NEAR/2

bird) OR (non-indigenous NEAR/2 crab) OR (non-indigenous NEAR/2 pine) OR (non-indigenous NEAR/2 vegetation) OR (non-indigenous NEAR/2 animal) OR (non-indigenous NEAR/2 goby) OR (non-indigenous NEAR/2 mosquito) OR (non-indigenous NEAR/2 parasitoid) OR (non-indigenous NEAR/2 rat) OR (non-indigenous NEAR/2 acacia) OR (non-indigenous NEAR/2 lionfish) OR (non-indigenous NEAR/2 amphipod) OR (non-indigenous NEAR/2 seaweed) OR (non-indigenous NEAR/2 grassland) OR (non-indigenous NEAR/2 parasite) OR (non-indigenous NEAR/2 carp) OR (non-indigenous NEAR/2 seed) OR (non-indigenous NEAR/2 herb) OR (non-indigenous NEAR/2 invertebrate) OR (non-indigenous NEAR/2 „bee“) OR (non-indigenous NEAR/2 rodent) OR (non-indigenous NEAR/2 vertebrate) OR (non-indigenous NEAR/2 salmonid) OR (non-indigenous NEAR/2 „vine“) OR (non-indigenous NEAR/2 congener) OR (non-indigenous NEAR/2 ascidian) OR (non-indigenous NEAR/2 flora) OR (non-indigenous NEAR/2 toad) OR (non-indigenous NEAR/2 spartina) OR (allochthonous NEAR/2 plant) OR (allochthonous NEAR/2 fish) OR (allochthonous NEAR/2 pest) OR (allochthonous NEAR/2 tree) OR (allochthonous NEAR/2 weed) OR (allochthonous NEAR/2 predator) OR (allochthonous NEAR/2 ant) OR (allochthonous NEAR/2 insect) OR (allochthonous NEAR/2 shrub) OR (allochthonous NEAR/2 crayfish) OR (allochthonous NEAR/2 mussel) OR (allochthonous NEAR/2 taxon) OR (allochthonous NEAR/2 pathogen) OR (allochthonous NEAR/2 plantation) OR (allochthonous NEAR/2 mammal) OR (allochthonous NEAR/2 earthworm ×) OR (allochthonous NEAR/2 forest) OR (allochthonous NEAR/2 algae) OR (allochthonous NEAR/2 trout) OR (allochthonous NEAR/2 herbivore) OR (allochthonous NEAR/2 snail) OR (allochthonous NEAR/2 beetle) OR (allochthonous NEAR/2 bivalve) OR (allochthonous NEAR/2 forb) OR (allochthonous NEAR/2 macrophyte) OR (allochthonous NEAR/2 bird) OR (allochthonous NEAR/2 crab) OR (allochthonous NEAR/2 pine) OR (allochthonous NEAR/2 vegetation) OR (allochthonous NEAR/2 animal) OR (allochthonous NEAR/2 goby) OR (allochthonous NEAR/2 mosquito) OR (allochthonous NEAR/2 parasitoid) OR (allochthonous NEAR/2 rat) OR (allochthonous NEAR/2 acacia) OR (allochthonous NEAR/2 lionfish) OR (allochthonous NEAR/2 amphipod) OR (allochthonous NEAR/2 seaweed) OR (allochthonous NEAR/2 grassland) OR (allochthonous NEAR/2 parasite) OR (allochthonous NEAR/2 carp) OR (allochthonous NEAR/2 seed) OR (allochthonous NEAR/2 herb) OR (allochthonous NEAR/2 invertebrate) OR (allochthonous NEAR/2 „bee“) OR (allochthonous NEAR/2 rodent) OR (allochthonous NEAR/2 vertebrate) OR (allochthonous NEAR/2 salmonid) OR (allochthonous NEAR/2 „vine“) OR (allochthonous NEAR/2 congener) OR (allochthonous NEAR/2 ascidian) OR (allochthonous NEAR/2 flora) OR (allochthonous NEAR/2 toad) OR (allochthonous NEAR/2 spartina) OR (exotic NEAR/2 plant) OR (exotic NEAR/2 fish) OR (exotic NEAR/2 pest) OR (exotic NEAR/2 tree) OR (exotic NEAR/2 weed) OR (exotic NEAR/2 predator) OR (exotic NEAR/2 ant) OR (exotic NEAR/2 insect) OR (exotic NEAR/2 shrub) OR (exotic NEAR/2 crayfish) OR (exotic NEAR/2 mussel) OR (exotic NEAR/2 taxon) OR (exotic NEAR/2 pathogen) OR (exotic NEAR/2 plantation) OR (exotic NEAR/2 mammal) OR (exotic NEAR/2 earthworm) OR (exotic NEAR/2 forest) OR (exotic NEAR/2 algae) OR (exotic NEAR/2 trout) OR (exotic NEAR/2 herbivore) OR (exotic NEAR/2 snail) OR (exotic NEAR/2 beetle) OR (exotic NEAR/2 bivalve ×) OR (exotic NEAR/2 forb) OR (exotic NEAR/2 macrophyte) OR (exotic NEAR/2 bird) OR (exotic NEAR/2 crab) OR (exotic NEAR/2 pine) OR (exotic NEAR/2 vegetation) OR (exotic NEAR/2 animal) OR (exotic NEAR/2 goby) OR (exotic NEAR/2 mosquito) OR (exotic NEAR/2 parasitoid ×) OR (exotic NEAR/2 rat) OR (exotic NEAR/2 acacia) OR (exotic NEAR/2 lionfish) OR (exotic NEAR/2 amphipod) OR (exotic NEAR/2 seaweed) OR (exotic NEAR/2 grassland) OR (exotic NEAR/2 parasite) OR (exotic NEAR/2 carp) OR (exotic NEAR/2 seed) OR (exotic NEAR/2 herb) OR (exotic NEAR/2 invertebrate) OR (exotic NEAR/2 „bee“) OR (exotic NEAR/2 rodent) OR (exotic NEAR/2 vertebrate) OR (exotic NEAR/2 salmonid) OR (exotic NEAR/2 „vine“) OR (exotic NEAR/2 congener) OR (exotic NEAR/2 ascidian) OR (exotic NEAR/2 flora) OR (exotic NEAR/2 toad) OR (exotic NEAR/2 spartina)))) AND DOCUMENT TYPES: (Article OR Review)

A.2. Übersicht Sample Invasionsbiologie

A.2.1. Phase 1: Analyse Klassifikation von Zitationsfunktionen

anwendungsorientiert_1_2159

Carson, B. D., Lishawa, S. C., Tuchman, N. C., Monks, A. M., Lawrence, B. A., & Albert, D. A. (2018). Harvesting invasive plants to reduce nutrient loads and produce bioenergy: an assessment of Great Lakes coastal wetlands. *Ecosphere*, 9(6). doi:10.1002/ecs2.2320

Inhaltsangabe: In den Küstenfeuchtgebieten der Great Lakes von Laurentiana bedrohen invasive Pflanzenarten das heimische Ökosystem. In der Studie geht es darum (1) das Ausmaß der Verbreitung der Arten abzuschätzen, (2) deren Biomasse zu ermitteln (nitrogen- und phosphorus-Gehalt) um (3) das Potential der Pflanzen für die (erneuerbare) Energiegewinnung zu bewerten. Insbesondere die Methode des mechanischen Entfernens (Ernte/harvesting) wird hinsichtlich ihrer Wirtschaftlichkeit und positiven Effekten auf das Ökosystem evaluiert.

research problem: Bewertung der (ökonomischen) Potentiale von Maßnahmen (insbesondere das mechanische Entfernen/Ernte/harvesting) gegen die Ausbreitung invasiver Schilffarten

empirical object: Eigenschaften und Habitat (Great Lakes, GLCW) der drei invasiven Pflanzen "common reed" (*Phragmites australis* (Cav.) Steud.), "reed canary grass" (*Phalaris arundinacea* L.) und invasive cattail" (*Typha x glauca*)

method: Feldstudien und Vergleiche schon vorhandener Pflanzensurveys (Literatur), Bio-Energie-Analysen

theoretical: Bedrohung der Biodiversität, Zusammenhänge zwischen Umweltvariablen und Invasionsverhalten/Verbreitung

application: negative Folgen durch die Verbreitung der Arten (Fischsterben, Rückgang des Tourismus)

anwendungsorientiert_2_1287

Kocovsky, P., Sturtevant, R., & Schardt, J. (2018). What it is to be established: policy and management implications for non-native and invasive species. *Management of Biological Invasions*, 9(3), 177–185. doi:10.3391/mbi.2018.9.3.01

Inhaltsangabe: Es geht darum zu klären, wie die Verwendung bestimmter Terminologien (konkret: der Begriff "established") Managemententscheidungen beeinflusst.

research problem: Überbrückung der Kluft zwischen ökologischer Forschung (Theorie) und dem Management invasiver Arten (Praxis)

empirical object: verschiedene

method: konzeptionell

theoretical:

application:

anwendungsorientiert_3_573

Stringham, O. C., & Lockwood, J. L. (2018). Pet problems: Biological and economic factors that influence the release of alien reptiles and amphibians by pet owners. *Journal of Applied Ecology*, 55(6), 2632–2640. doi:10.1111/1365-2664.13237

Inhaltsangabe: Der Handel mit exotischen Haustieren (Reptilien und Amphibien) hat großen Anteil an der Einführung neuer (invasiver) Arten in lokale Ökosysteme. Die Studie untersucht wann die Wahrscheinlichkeit, dass Haustierbesitzer ihre Tiere in die Wildnis aussetzen am höchsten ist, bzw. von welchen Faktoren die Aussetz-Entscheidung abhängt. Dazu werden Sekundärdaten zu biologischen und ökonomischen Eigenschaften verschiedener exotischer Haustierarten analysiert.

research problem: (biologische und ökonomische) Faktoren, die die Aussetz-Entscheidung von Haltern exotischer Haustiere beeinflussen

empirical object: Reptilien- und Amphibien-Arten und deren Eigenschaften

method: statistische Analyse (Regressionsmodelle)

theoretical: generelle Hypothesen und Theorien zum research problem, faktoren, die die durchsetzungsfähigkeit der invasiven arten bestimmen

application: Policy Implications, Verhinderung von Aussetzung durch Information und Aufklärung unterstützen, negative auswirkungen

anwendungsorientiert_4_4261

Stien, J., & Hausner, V. H. (2018). Motivating and engaging volunteer hunters to control the invasive alien American mink *Neovison vison* in Norway. *Oryx - The International Journal of Conservation*, 52(1), 186–194. doi:10.1017/S0030605316000879

Inhaltsangabe: Es wurde eine Befragung unter den Mitgliedern der norwegischen Jäger- und Anglerverbandes durchgeführt um deren Einstellungen zu einem eventuellen Nerz-Kontrollprogramm abzufragen. Hintergrund: Der Nerz als invasive Art scheint in Norwegen ein Problem zu sein und es sollen Lösungswege entwickelt werden die Verbreitung der Art zu begrenzen.

research problem: Interesse an Nerz Kontrollprogrammen, Motivation für Nerz Eindämmung

empirical object: Nerz, amerikan mink (neovison vison) in Nord-Europa, Heimatgebiet Amerika, Invasionsgebiet Nord-Europa

method: Umfrage bei den Mitgliedern des norwegischen Jäger- und Angler-Verbandes

theoretical: allgemeine Invasionstheorie

application: Management-Maßnahmen zur Eindämmung der Nerz Population (z.B. monetary rewards, social benefits, etc.)

empirisch_1_2273

Lopez-Gappa, J., & Liuzzi, M. G. (2018). Recent discovery of non-indigenous bryozoans in the fouling assemblage of Quequen Harbour (Argentina, Southwest Atlantic). *Marine Biodiversity*, 48(2), 1159–1167. doi:10.1007/s1252601605617

Inhaltsangabe: Es geht um zwei verschiedene Arten von Moostierchen (bryozoans), die das erste Mal in Argentinien gesichtet wurden. Sie wurden auf experimentellen Substraten gefunden, die zur Untersuchung der Bewuchsansammlungen in Quequen Harbour ausgebracht waren.

Bemerkung: einige Quellen sind auf Spanisch, Italienisch oder Französisch, alt (viel 1980er Jahre, aber auch 1820er Jahre) und nicht in WoS indiziert, was die Zuordnung erschwert, da Abstracts oder Volltexte für diese Arbeiten oft schwer zu finden sind. Bei einigen der sehr alten Quellen handelt es sich um Atlanten, in denen verschiedene Arten (das erste Mal) beschrieben und bebildert werden.?

research problem: Verteilung/Verbreitung/Anwesenheit von invasiven Moostierchen (in Häfen weltweit)

empirical object: die zwei invasiven Moostierchen-Arten *Aspidelectra Melolontha*, *Amathia Verticillata* in Quequen Harbour und deren Eigenschaften und Lebensraum (Habitat)

method: Feldbeobachtungen und genetische Analyse der Tierchen

theoretical: Verbreitungswege (Schifffahrt)

application: Kontrolle der Ausbreitung, Risikobewertung

empirisch_2_2845

Wellband, K. W., Pettitt-Wade, H., Fisk, A. T., & Heath, D. D. (2018). Standing genetic diversity and selection at functional gene loci are associated with differential invasion success in two nonnative fish species. *Molecular ecology*, 27(7), 1572–1585. doi:10.1111/mec.14557

Inhaltsangabe: Zwei Fischarten wurden auf ihren Invasionserfolg getestet. Dabei wurden neue funktionale genetische Marker verwendet.

research problem: genetische Eigenschaften, die den Invasionserfolg beeinflussen, insbesondere heterozygosity

empirical object: die zwei invasiven Fisch-Arten Tubenose Goby, Round Gobi in den Great Lakes und deren Eigenschaften und Lebensraum (Habitat)

method: Vergleichende Genanalysen beider Arten mittels single nucleotide polymorphisms (SNPs) und microsattelites

theoretical: theoretische Annahmen, Hypothesen und Definitionen zu Invasion, Invasionserfolg, Evolutionsprozessen und solche die in Modelle einfließen, „genetic drift effect“, „natural selection“, „genetic bottlenecks“, „gene surfing“

application: Management der invasiven Art, Einschätzung und Eindämmung der Risiken

empirisch_3_2754

Nagy, D. U., Stranzinger, S., Godi, A., Weisz, A., Rosche, C., Suda, J., Mariano, M. & Pal, R. W. (2018). Does higher ploidy level increase the risk of invasion? A case study with two geocytotypes of *Solidago gigantea* Aiton (Asteraceae). *Journal of Plant Ecology*, 11(2), 317–327. doi:10.1093/jperts005

Inhaltsangabe: In der Studie wurden Pflanzen der riesen Goldrute mit unterschiedlichen Chromosomenanzahlen verglichen. Es konnte die Hypothese widerlegt werden, dass sich ein höheres ploidy level positiv auf den Invasionserfolg auswirkt. Die hexaploiden Pflanzen aus dem Heimatgebiet Montana konnten sich weniger gut an veränderte Umweltbedingungen im Common Garden anpassen, als tetraploide Pflanzen, die bereits in Europa invasiv geworden sind.

research problem: Einfluss genetischer Eigenschaften (insbesondere ploidy level, genetic diversity, cytogeography) auf den Invasionserfolg?

empirical object: *Solidago gigantea* (riesen Goldruten, giant goldenrod, Asteraceae) im Invasionsgebiet Europa und Heimatgebiet Nordamerika

method: greenhouse and common garden experiments, along with microsatellite analyses
theoretical: allgemeine Aussagen zu invasiven Arten, Invasionsprozessen und ihren biogeographischen Ursprüngen, EICA hypothesis
application: Management (zukünftiger) invasiver Arten

modellierung_1_1509

Singh, K. K., Chen, Y.-H., Smart, L., Gray, J., & Meentemeyer, R. K. (2018). Intraannual phenology for detecting understory plant invasion in urban forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, 142, 151–161. doi:10.1016/j.isprsjprs.2018.05.023

Inhaltsangabe: In der städtischen Waldregion um Mecklenburg, USA geht um die Erkennung einer speziellen invasiven Pflanzenart (*Linustrum Sinese*) im Unterwuchs in bewaldeten Umgebungen. In der Studie wurden Fernerkundungsbilder genutzt und damit Modelle erstellt um die Verbreitung dieser invasiven Art zu berechnen. Neben den selbst erhobenen Felddaten nutzen die Autoren noch verschiedene andere Datenquellen und -arten wie Luftaufnahmen, Geodaten und hydrografische Daten.

research problem: Erkennung der Verbreitung invasiver Pflanzen (understory invasive species detection), auch allgemeine Bemerkungen zu Modellierungsansätzen

empirical object: Eigenschaften und Habitat der invasiven Pflanzenart *Linustrum Sinese* im Unterwuchs in bewaldeten Gebieten in North Carolina

method: Fernbilder durch Landsat TM, Feldbeobachtungen, Luftaufnahmen und Modellierung anhand von Vegetationsindizes (VI), Topographieindizes (TI) und einem Random Forest Model (RF), häufig spezifische Nennung einzelner Indizes (SRI, TRI, TMI, ...), Formeln (NVDI, EVI, MAVI, ...) und Modelle, konkretes Vorgehen in der vorliegenden Studie unter Verweis auf Modelle usw. aus anderen Arbeiten

theoretical: keine

b Konsequenzen und Gefahren für Biodiversität und Menschen durch die Verbreitung, z.B. auch für Relevanzbegründung

modellierung_2_73

Shrestha, U. B., Sharma, K. P., Devkota, A., Siwakoti, M., & Shrestha, B. B. (2018). Potential impact of climate change on the distribution of six invasive alien plants in Nepal. *Ecological Indicators*, 95, 99–107. doi:10.1016/j.ecolind.2018.07.009

Inhaltsangabe: Auf Grundlage von vorgehender Feldstudien und unter Berücksichtigung verschiedener klimatischer Szenarien wird die Ausbreitung der sechs invasivsten und problematischsten Pflanzenarten in Nepals Bergregionen prognostiziert.

research problem: Verbreitung invasiver Pflanzenarten unter Berücksichtigung verschiedener klimatischer Szenarien und deren Prognose

empirical object: sechs der problematischsten invasiven Pflanzenarten in Nepal (*Ageratum houstonianum* Mill., *Chromolaena odorata* (L.), *Hyptis suaveolens* (L.), *Lantana camara* L., *Mikania micrantha*, and *Parthenium hysterophorus* L.)

method: Modellierung auf Basis von Fundorten (aus eigenen field surveys), Umwelt- und Klimavariablen (aus Sekundärdaten), häufig spezifische Nennung einzelner Modelle, Formeln und Indizes

theoretical: generelle Trends in Invasionsprozessen im Zusammenhang mit dem Klimawandel

application: Nationale und politische Programme zur Eindämmung der invasiven Arten, generell Management

modellierung_3_4256

Jazwa, M., Jedrzejczak, E., Klichowaska, E., & Pliszko, A. (2018). Predicting the potential distribution area of *Solidago x niedereideri* (Asteraceae). *Turkish Journal of Botany*, 42, 51–56. doi:10.3906/bot1703-17

Inhaltsangabe: Es wurde die potentielle Verbreitung einer hybriden Art der riesen Goldrute mittels eines Modellierungsansatzes (MAXENT) prognostiziert. ?Auf Grundlage von Fundort-Daten und unter Berücksichtigung verschiedener Umweltvariablen wird die weitere Ausbreitung der invasiven riesen Goldrute in Europa mit dem Modellierungsansatz MAXENT prognostiziert.

research problem: Verbreitung hybrider invasiver Arten unter Berücksichtigung von Umweltvariablen und deren Prognose

empirical object: Eigenschaften und Habitat der *Solidago x niedereideri* (Asteraceae) und deren Elternpflanzen aus denen sich das Hybrid entwickelt hat, Angaben zu den Fundorten der Elternpflanzen werden hier immer als emp obj kodiert, weil es nicht um das rp geht (rp = Verbreitung hybrider invasiver Arten)

method: Modellierungsansatz MAXENT

theoretical: allgemeine Theorien der Invasionsbiologie

application: Negative Einflüsse der Pflanzenart und Management

A.2.2. Phase 2: Netzwerkanalyse

sd

van Riel, M. C., van der Velde, G., & bij de Vaate, A. (2009). Interference competition between alien invasive gammaridean species. *Biological Invasions*, 11(9), 2119–2132. doi:10.1007/s10530-009-9486-4

r1

Leuven, R. S. E. W., van der Velde, G., Baijens, I., Snijders, J., van der Zwart, C., Lenders, H. J. R., & bij de Vaate, A. (2009). The river Rhine: a global highway for dispersal of aquatic invasive species. *Biological Invasions*, 11(9), 1989–2008. doi:10.1007/s10530-009-9491-7

r2

van Kleef, H., van der Velde, G., Leuven, R. S. E. W., & Esselink, H. (2008). Pumpkinseed sunfish (*Lepomis gibbosus*) invasions facilitated by introductions and nature management strongly reduce macroinvertebrate abundance in isolated water bodies. *Biological Invasions*, 10(8), 1481–1490. doi:10.1007/s10530-008-9220-7

r3

van Riel, M. C., van der Velde, G., Rajagopal, S., Marguillier, S., Dehairs, F., & Vaate, A. b. de. (2006). Trophic Relationships in the Rhine Food Web during Invasion and after Establishment of the Ponto-Caspian Invader *Dikerogammarus villosus*. *Hydrobiologia*, 565(1), 39–58. doi:10.1007s10750-005-1904-8

r4

Platvoet, D., Dick, J. T.A., Konijnendijk, N., & van der Velde, G. (2006). Feeding on micro-algae in the invasive Ponto-Caspian amphipod *Dikerogammarus villosus* (Sowinsky, 1894). *Aquatic Ecology*, 40(2), 237–245. doi:10.1007s10452-005-9028-9

r5

Meyer, A., Kaschek, N., & Meyer, E. I. (2004). The effect of low flow and stream drying on the distribution and relative abundance of the alien amphipod, *echnogammarus berillioni* (Catta, 1878) in karstic stream system (Westphalia, Germany). *Crustaceana*, 77(8), 909–922.

r6

Bollache, L., Devin, S., Wattier, R., Chovet, M., Beisel, J.-N., Moreteau, J.-C., & Rigaud, T. (2004). Rapid range extension of the Ponto-Caspian amphipod *Dikerogammarus villosus* in France: potential consequences. *Archiv für Hydrobiologie*, 160(1), 57–66. doi:10.11270003-913620040160-0057

r7

Jazdzewski, K., Konopacka, A., & Grabowski, M. (2004). Recent drastic changes in the gammarid fauna (Crustacea, Amphipoda) of the Vistula River deltaic system in Poland caused by alien invaders. *Diversity and Distributions*, 10, 81–87.

r8

Wijnhoven, S., van Riel, M. C., & van der Velde, G. (2003). Exotic and indigenous freshwater gammarid species: physiological tolerance to water temperature in relation to ionic content of the water. *Aquatic Ecology*, 37, 151–158.

r9

van Overdijk, Colin D. A., Grigorovich, I. A., Mabee, T., Ray, W. J., Ciborowski, J. J. H., & MacIsaac, H. J. (2003). Microhabitat selection by the invasive amphipod *Echinogammarus ischnus* and native *Gammarus fasciatus* in laboratory experiments and in Lake Erie. *Freshwater Biology*, 48, 567–578.

r10

MacNeil, C., Bigsby, E., Dick, J. T. A., Hatcher, M. J., & Dunn, A. M. (2003). Differential physico-chemical tolerances and intraguild predation among native and invasive amphipods (Crustacea); a field study. *Archiv für Hydrobiologie*, 156(2), 165–179. doi:10.11270003-913620030156-0165

r11

Dick, J. T.A., Platvoet, D., & Kelly, D. W. (2002). Predatory impact of the freshwater invader *Dikerogammarus villosus* (Crustacea: Amphipoda). *Canadian Journal of Fisheries and Aquatic Sciences*, 59(6), 1078–1084. doi:10.1139F02-074

r12

Dick, J. T. A., & Platvoet, D. (2000). Invading predatory crustacean *Dikerogammarus villosus* eliminates both native and exotic species. *Proceedings of the Royal Society B. Biological Science*, 267, 977–983.

z1

Penk, M. R., Donohue, I., & Irvine, K. (2018). Temporally variable niche overlap and competitive potential of an introduced and a native mysid shrimp. *Hydrobiologia*, 823(1), 109–119. doi:10.1007s10750-018-3700-2

z2

Clinton, K. E., Mathers, K. L., Constable, D., Gerrard, C., & Wood, P. J. (2018). Substrate preferences of coexisting invasive amphipods, *Dikerogammarus villosus* and *Dikerogammarus haemobaphes*, under field and laboratory conditions. *Biological Invasions*, 20(8), 2187–2196. doi:10.1007s10530-018-1695-2

z3

Koester, M., Schneider, M., Hellmann, C., Becker, J., Winkelmann, C., & Gergs, R. (2018). Is The invasive amphipod *Dikerogammarus villosus* the main factor structuring the benthic community across different types of water bodies in the River Rhine system? *Limnologia*, 71, 44–50. doi:10.1016j.limno.2018.06.001

z4

Alther, R., & Altermatt, F. (2018). Fluvial network topology shapes communities of native and non-native amphipods. *Ecosphere*, 9(2), e02102.

z5

Mauchart, P., Czirik, A., Horvai, V., Herczeg, R., Móra, A., & Csabai, Z. (2017). Effects of meso- and microhabitat characteristics on the coexistence of two native gammarid species (Crustacea, Gammaridae). *International Review of Hydrobiology*, 102(1-2), 38–46. doi:10.1002/iroh.201601855

z6

Kobak, J., Jermacz, ?, Rutkowska, D., Pawłowska, K., Witkowska, L., & Poznańska, M. (2017). Impact of predators and competitors on the depth selection by two invasive gammarids. *Journal of Zoology*, 301(3), 174–183. doi:10.1111/jzo.12409

z7

Gelder, S. de, van der Velde, G., Platvoet, D., Leung, N., Dorenbosch, M., Hendriks, H.W.M., & Leuven, R.S.E.W. (2016). Competition for shelter sites: Testing a possible mechanism for gammarid species displacements. *Basic and Applied Ecology*, 17(5), 455–462. doi:10.1016/j.baae.2016.01.008

z8

Hellmann, C., Worischka, S., Mehler, E., Becker, J., Gergs, R., & Winkelmann, C. (2015). The trophic function of *Dikerogammarus villosus* (Sowinsky, 1894) in invaded rivers: a case study in the Elbe and Rhine. *Aquatic Invasions*, 10(4), 385–397. doi:10.3391/ai.2015.10.4.03

z9

Gergs, R., & Rothhaupt, K.-O. (2015). Invasive species as driving factors for the structure of benthic communities in Lake Constance, Germany. *Hydrobiologia*, 746(1), 245–254. doi:10.1007/s10750-014-1931-1934

z10

Rewicz, T., Wattier, R., Grabowski, M., Rigaud, T., & B?cela-Spychalska, K. (2015). Out of the Black Sea: phylogeography of the invasive killer shrimp *Dikerogammarus villosus* across Europe. *PloS one*, 10(2), e0118121. doi:10.1371/journal.pone.0118121

z11

Mauchart, P., Bereczki, C., Ortmann-Ajkai, A., Csabai, Z., & Szivák, I. (2014). Niche segregation between two closely similar gammarids (Peracarida, Amphipoda) - native vs. naturalized non-native species. *Crustaceana*, 87(11-12), 1296–1314. doi:10.1163/15685403-00003355

z12

Rewicz, T., Grabowski, M., MacNeil, C., & B?cela-Spychalska, K. (2014). The profile of a „perfect“ invader - the case of killer shrimp, *Dikerogammarus villosus*. *Aquatic Invasions*, 9(3), 267–288. doi:10.3391/ai.2014.9.3.04

z13

Russell, J. C., Sataruddin, N. S., & Heard Allison D. (2014). Over-invasion by functionally equivalent invasive species. *Ecology*, 95(8), 2268–2276.

z14

Gergs, R., Schlag, L., & Rothhaupt, K.-O. (2013). Different ammonia tolerances may facilitate spatial coexistence of *Gammarus roeselii* and the strong invader *Dikerogammarus villosus*. *Biological Invasions*, 15(8), 1783–1793. doi:10.1007/s10530-013-0408-0

z15

Oyugi, D. O., Cucherousset, J., & Robert Britton, J. (2012). Temperature-dependent feeding interactions between two invasive fishes competing through interference and exploitation. *Reviews in Fish Biology and Fisheries*, 22(2), 499–508. doi:10.1007/s11160-011-9243-5

z16

Hänfling, B., Edwards, F., & Gherardi, F. (2011). Invasive alien Crustacea: dispersal, establishment, impact and control. *BioControl*, 56(4), 573–595. doi:10.1007/s10526-011-9380-8

z17

Hudina, S., Gali?, N., Roessink, I., & Hock, K. (2011). Competitive interactions between co-occurring invaders: identifying asymmetries between two invasive crayfish species. *Biological Invasions*, 13(8), 1791–1803. doi:10.1007/s10530-010-9933-2

z18

Lagrange, C., Kaldonski, N., Motreuil, S., Lefèvre, T., Blatter, O., Giraud, P., & Bollache, L. (2011). Interspecific differences in drift behaviour between the native *Gammarus pulex* and the exotic *Gammarus roeseli* and possible implications for the invader’s success. *Biological Invasions*, 13(6), 140–1421. doi:10.1007/s10530-010-9899-0

z19

PISCART, C., ROUSSEL, J.-M., Dick, J. T. A., GROSBOS, G., & MARMONIER, P. (2011). Effects of coexistence on habitat use and trophic ecology of interacting native and invasive amphipods. *Freshwater Biology*, 56(2), 325–334. doi:10.1111/j.1365-2427.2010.02500.x

z20

Stoffels, G. E. M. W., Tummers, J. S., van der Velde, G., Platvoet, D., Hendriks, H.W.M., & Leuven, R. S. E. W. (2011). Assessment of predatory ability of native and non-native freshwater gammaridean species: A rapid test with water fleas as prey. *Current Zoology*, 57(6), 836–843.

z21

Leuven, R. S. E. W., van der Velde, G., Baijens, I., Snijders, J., van der Zwart, C., Lenders, H. J. R., & bij de Vaate, A. (2009). The river Rhine: a global highway for dispersal of aquatic invasive species. *Biological Invasions*, 11(9), 1989–2008. doi:10.1007/s10530-009-9491-7

z22

van der Velde, G., Leuven, R. S. E. W., Platvoet, D., Bacela, K., Huijbregts, M. A. J., Hendriks, H. W. M., & Kruijt, D. (2009). Environmental and morphological factors influencing predatory behaviour by invasive non-indigenous gammaridean species. *Biological Invasions*, 11(9), 2043–2054. doi:10.1007/s10530-009-9500-x

B. Anhang – weitere Tabellen und Abbildungen zur Auswertung der Ergebnisse

B.1. Tabellen

Absolute und erwartete Häufigkeiten von epistemic function nach Dokumentgruppe

```
> addmargins(table(dataset_ref$epistemic_function, dataset_ref$Dokumentgruppe))
```

	anwendungsorientiert	empirisch	modellierung	Sum
application	49	5	28	82
empirical object	68	134	37	239
method	42	109	113	264
research problem	65	121	79	265
theoretical	40	48	21	109
Sum	264	417	278	959

```
> addmargins(chi2test_1$expected)
```

	anwendungsorientiert	empirisch	modellierung	Sum
application	22.57351	35.65589	23.77059	82
empirical object	65.79353	103.92388	69.28259	239
method	72.67570	114.79458	76.52972	264
research problem	72.95099	115.22941	76.81960	265
theoretical	30.00626	47.39625	31.59750	109
Sum	264.00000	417.00000	278.00000	959